Regulation and implementation of usability engineering for a medical device

Melissa Holopainen

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.
Helsinki 21.02.2020

Supervisor

Prof. Risto Ilmoniemi

Advisor

Markus Paasi, M.Sc (Tech)

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Melissa Holopainen

Title Regulation and implementation of usability engineering for a medical device

Degree programme Life Science Technologies

Major Biomedical Engineering **Code of major** SCI3059

Supervisor Prof. Risto Ilmoniemi

Advisor Markus Paasi, M.Sc (Tech)

Date 21.02.2020 **Number of pages** 53+5 **Language** English

Abstract

The safety and effectiveness of medical devices are ensured by market-area-specific laws and regulations. In 2017, the European Parliament and Council published a new medical device regulation (MDR) that repealed the old medical device directives. Considering usability, the regulation states that the device manufacturer must eliminate or reduce the risks related to use errors as far as possible. Complex user interfaces (UI) that have inadequate usability can cause use errors that can possibly affect the safety of a patient or a user.

The goal of this master's thesis was to research medical device regulation, update the quality system of a company to comply with the new MDR considering usability, develop a usability engineering (UE) process for the company according to usability standard IEC 62366-1:2015 and apply the process for a new navigated transcranial magnetic stimulation (nTMS) device.

The quality management system (QMS) of the company was updated by developing a new version of a standard operating procedure (SOP) for the UE process. The UE process resulted in finding several use errors and hazards that could affect the safety of a patient or a user. The usability evaluation performed for the new nTMS device resulted in the identification of 17 usability findings that were mostly created by unclear instruction texts and ambiguous GUI components.

Keywords Medical Device Regulation, medical device, graphical user interface, usability, usability engineering process, IEC 62366-1:2015

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Diplomityön tiivistelmä

Tekijä Melissa Holopainen

Työn nimi Asetukset ja käytettävyystekniikan soveltaminen lääkinnälliselle laitteelle

Koulutusohjelma Life Science Technologies

Pääaine Biomedical Engineering **Pääaineen koodi** SCI3059

Työn valvoja Prof. Risto Ilmoniemi

Työn ohjaaja Markus Paasi, M.Sc (Tech)

Päivämäärä 21.02.2020 **Sivumäärä** 53+5 **Kieli** Englanti

Tiivistelmä

Lääkinnällisten laitteiden turvallisuus ja tehokkuus turvataan markkina-aluekohtaisilla laeilla ja asetuksilla. Vuonna 2017, Euroopan parlamentti ja neuvosto julkaisivat uuden asetuksen lääkinnällisistä laitteista kumoten aikaisemmat lääkinnällisten laitteiden direktiivit. Käytettävyydestä asetus määrää, että laitteiden valmistajan on poistettava tai vähennettävä käyttövirheistä johtuvia riskejä niin paljon kuin mahdollista. Monimutkaiset käyttöliittymät puutteellisella käytettävyydellä voivat aiheuttaa käyttövirheitä, jotka voivat vaarantaa potilaan tai käyttäjän turvallisuuden.

Tämän diplomityön tavoite oli tutkia lääkinnällisten laitteiden sääntelyä, päivittää firman laatujärjestelmä lääkinnällisten laitteiden asetuksen mukaiseksi käytettävyyden osalta, kehittää käytettävyystekniikkaprosessi firmalle käytettävyysstandardi IEC 62366-1:2015 mukaisesti ja soveltaa prosessia uuteen navigoituun transkraniaaliseen magneettistimulaatiolaitteeseen (nTMS laite).

Firman laatujärjestelmä päivitettiin laatimalla uusi versio käytettävyystekniikkaprosessin toimintaohjeesta. Käytettävyystekniikkaprosessissa löydettiin useita mahdollisia käyttövirheitä ja vaaroja, jotka voisivat vaikuttaa potilaan tai käyttäjän turvallisuuteen. Uudella nTMS laitteella suoritetussa käytettävyysarvioinnissa tunnistettiin 17 käytettävyyslöydöstä, jotka johtuivat enimmäkseen epäselvistä ohjeteksteistä ja käyttöliittymäkomponenteista.

Avainsanat Lääkinnällisten laitteiden asetus, lääkinnällinen laite, graafinen käyttöliittymä, käytettävyys, käytettävyystekniikkaprosessi, IEC 62355-1:2015

Preface

I want to express my gratitude to professor Risto Ilmoniemi for supervising my thesis and having time to comment on my writing. I want to thank my advisor Markus Paasi for his excellent guidance and help throughout the thesis. A big thank you also to my awesome colleagues at Nexstim for answering my questions and bringing joy to my work days. I would also like to thank all the clinical professionals who participated in my usability study.

Studying at university would have been a lot harder without my dear friends and fellow students. Thank you for your help with long days of studying and all the joy you brought me after school. Finally, I would like to express my endless gratitude to my mom and dad for their support and encouragement throughout my studies.

Helsinki, 21.02.2020

Melissa Holopainen

Contents

Abbreviations

3D	Three dimensional
AAMI	Association for the Advancement of Medical Instrumentation
AMID	Active Implantable Medical Devices Directive
CE	Conformité Européenne
DHF	Design history file
DHR	Device history record
DLPFC	Dorsolateral prefrontal cortex
DMR	Device master record
E-field	Electric field
EU	European Union
EMG	Electromyography
FDA	Food and Drug Administration
GUI	Graphical user interface
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
IVDD	In Vitro Diagnostic Directive
MCE	Moderate critical error
MDAA	Medical Device Amendments Act
MDD	Medical Device Directive
NCE	Non-critical error
MDR	Medical Device Regulation
MFC	Microsoft Foundation Class Library
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MT	Motor Threshold
NBT	Navigated Brain Therapy, product name for the company's therapeutic device
NBS	Navigated Brain Stimulation, product name for the company's diagnostic device
nTMS	Navigated TMS
POI	Point of interest
PMA	Pre-market approval
PMS	Post-market surveillance
QA&RA	Quality Assurance and Regulatory Affairs
QML	Qt Modeling Language
QMS	Quality management system
rTMS	Repetitive TMS
SNBT	Smart Navigated Brain Therapy, product name for the company's new device the UE process is applied for
SOP	Standard operating procedure
TMS	Transcranial magnetic stimulation
TPD	Therapeutic Products Directorate
UI	User interface
UE	Usability engineering
US	United States
WHO	World Health Organization

1 Introduction

Development of modern medical technology started in the first half of the 19th century and since then the number of devices has increased rapidly and they have become an essential part of health care [1]. For the first 100 years, there were no regulation and fraudulent and dangerous devices were marketed. In the beginning of the 20th century in the United States (US), the Pure Food and Drugs Act was signed into a law eventually leading to the creation of the Food and Drug Administration (FDA) and medical device regulation. [2] Similarly, other countries have established their own regulations and have assigned authorities to ensure the safety of patients and people using medical devices.

Medical devices are regulated to ensure the effectiveness and safety of such devices. In the US, the FDA is the regulatory authority that supervises that medical devices entering the US market obey the national laws. Before a new medical device can be released to the market, it needs to get approval from the supervising authorities. Before the 1990, all European countries had their own way of regulating medical devices [3]. However, in 1990s, the Council of the European Union introduced three directives to harmonize the complex regulatory systems in the European Union (EU) [4, 5, 6].

In 2017, a new regulation on medical devices was adopted in EU replacing existing directives. Along with other requirements, the medical device regulation states that medical devices shall be designed and manufactured in such way that they are safe, effective and suitable for their intended purpose. Risks related to medical device use errors should be eliminated or reduced as far as possible. [7]

Laws and regulations are often hard to comprehend and thus, several organizations have written standards that comply with the regulations. A harmonized standard is a standard that is recognized by regulatory authorities to provide conformity to regulations. It is not obligatory to follow standards, but they can be used to prove the requirements of regulations. Standards often list the deliverable information a manufacturer needs to provide. The obligation of the usability requirements of the new medical device regulation (MDR) can be proven by applying two harmonized standards recognized by the EU and the FDA.

Nowadays, medical devices are used not only in hospitals but also in private practices and at home. As the use environment changes, also the users vary from different health care professionals to patients themselves. The usability of a medical device is an important factor to ensure correct use and safety of the user and the one receiving the treatment. Complex medical devices that have inadequate usability can cause use errors that possibly lead to dangerous situations. [8]

Navigated transcranial magnetic stimulation (nTMS) is an advanced technology used, for example, for presurgical mapping of patient's brain areas and treatment of depression. Navigated TMS devices have traditionally been quite complex to use and hard to learn. The company for which this master's thesis is conducted for is developing a next-generation nTMS device that will emphasize usability.

The purpose of this master's thesis is to

- research the medical device regulation in different market areas
- update the quality system of the company developing and manufacturing nTMS devices to the level of the new MDR considering usability requirements
- develop a usability engineering (UE) process for the company according to the two above mentioned harmonized standards
- apply the UE process for the company's new nTMS device

The standards cover the usability of the whole user interface including the physical aspects of the medical device as well as the software. As software is the new component of the nTMS device under development, the focus of this master's thesis is on the usability of the software although the UE process is applied to the whole user interface.

In this master's thesis, the background of the company, TMS and the device the usability engineering process is applied to are given in Section 2. The medical device regulatory systems of the most significant market areas of the device are presented in Section 3. In Section 4, a detailed description of the applied usability standards is given. Section 5 describes the quality system of the firm and how it was updated to comply with the new MDR. Results of the UE process are gone through in Section 6 and Section 7 emphasizes the usability evaluation performed in the process by presenting detailed description of the methods and obtained results. Further development ideas of the device and future tasks are discussed in Section 8 and the thesis is summarized in Section 9.

2 Background

This chapter gives a background of the company the usability engineering process is applied to, the nTMS technology the firm is using and existing and new products of the company.

2.1 Company

The company for which the usability engineering process is applied is an incorporated medical device firm that is registered and based in Finland. Nexstim designs, manufactures, markets, sells, and services non-invasive navigated brain stimulation products of the central nervous systems for therapeutic and diagnostic purposes.

Nexstim Plc is a parent company of the group and they share headquarters in Helsinki, Finland. In addition to the parent company, the group has assigned sales and service responsibilities to the fully owned affiliates Nexstim Inc in the US and Nexstim GmbH in Germany. Nexstim Inc. acts as an importer and distributor in the US.

The enterprise is publicly listed on both Nasdaq First North Growth Market Finland and Nasdaq First North Growth Market Sweden following an initial public offering in November 2014. The company relies on the unique nTMS technology and has over 70 patents and pending patent applications. The key market areas are in Europe, the US and Canada.

2.2 Transcranial magnetic stimulation

Transcranial magnetic stimulation (TMS) is a non-invasive brain stimulation method introduced in 1985 by Barker and colleagues [9]. In TMS, a strong electric current is fed into a round or figure-eight-shaped coil that is placed over the patient's head. The coil generates a pulsed magnetic field that induces an electric field (E-field) inside the brain. [10] The E-field affects the transmembrane potential of the neurons and consequently, the voltage-sensitive ion channels in the nerve cells [11]. Opening of the ion channels may lead to depolarization of the neurons which in turn leads to action potentials. Action potentials are signals that propagate inside neurons activating other neurons or muscles. [10]

TMS can be navigated or non-navigated. In non-navigated TMS, the coil is positioned over THE patient's head based on the shape of the head surface and locations of the eyes, ears, and nose. The disadvantage of non-navigated TMS is the lack of knowledge of the stimulation location. In navigated TMS, different technologies are used in order to show the stimulation location on the brain. There are two types of navigated TMS: line-navigated and E-field-navigated. In line navigation, the stimulation coil is placed over the target area based on the patient's individual magnetic resonance (MR) images. It is assumed that the maximal E-field will be located on a line at the center of the coil perpendicular to the coil surface. However, line navigation can lead into inaccurate stimulation location, if the coil is not positioned optimally in respect of the skull. In E-field-navigated TMS, in

addition to the location information gained from the MR images, the patient's head geometry is also taken into account when calculating the E-field location and strength. [10] A visualization of the E-field and line-navigated TMS are presented in Figure 1.

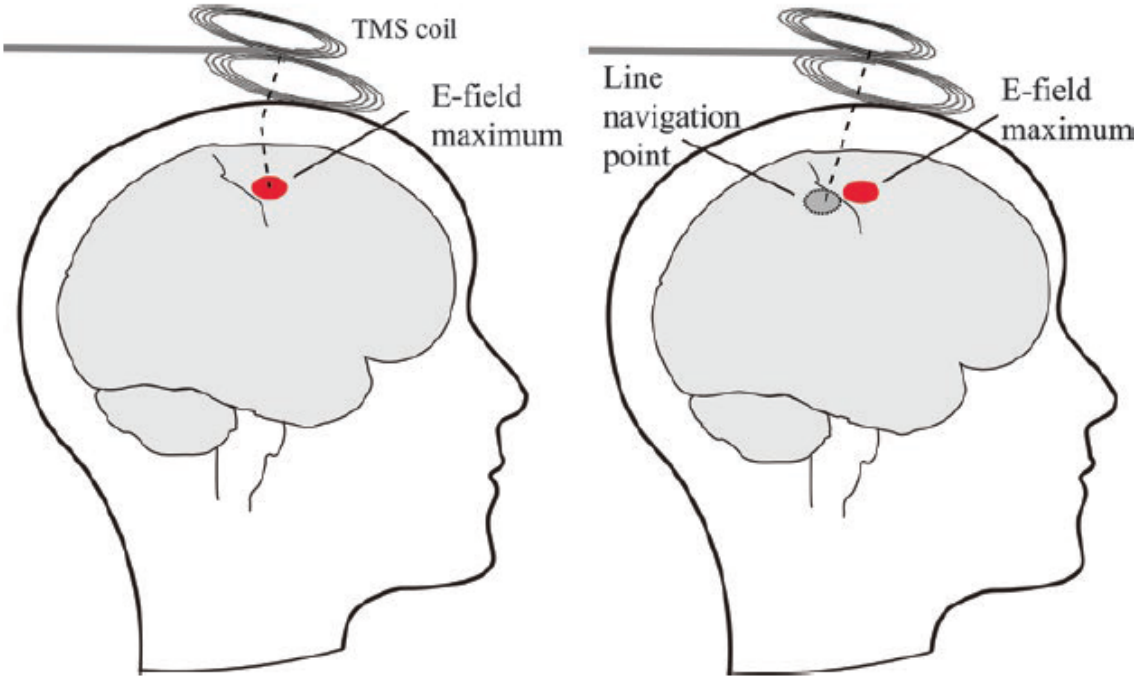


Figure 1: Difference between the navigation point location obtained using E-field navigation and line navigation. The image on the left side presents the E-field-navigated stimulation point and the E-field maximum at the same location. The image on the right side presents the line-navigated stimulation point and the actual location of the E-field maximum in slightly different locations. [10]

E-field navigation can be obtained when technical parameters of the stimulating coil and physical parameters of the patient's head and brain are known. Required coil parameters include location, orientation, tilt, size, and shape of the copper windings. Physical parameters of the patient's head can be obtained from an magnetic resonance image (MRI) scan that includes the whole head. Knowledge of the above parameters needs to be combined by performing a registration where the patient's head coordinates are aligned with the MRI's coordinates (Figure 2). When the head tracker and coil are in the camera's field of view, their locations are known in the camera coordinate system. Anatomical landmarks in relation to the head tracker are marked with a registration pen to the camera coordinate system. When the same landmarks are marked to the MRI scans, a registration algorithm can combine the MRI and camera coordinate systems. After the registration is performed, measuring the coil location in relation to the head tracker enables showing E-field location in the 3D reconstruction of the patient's MR images.

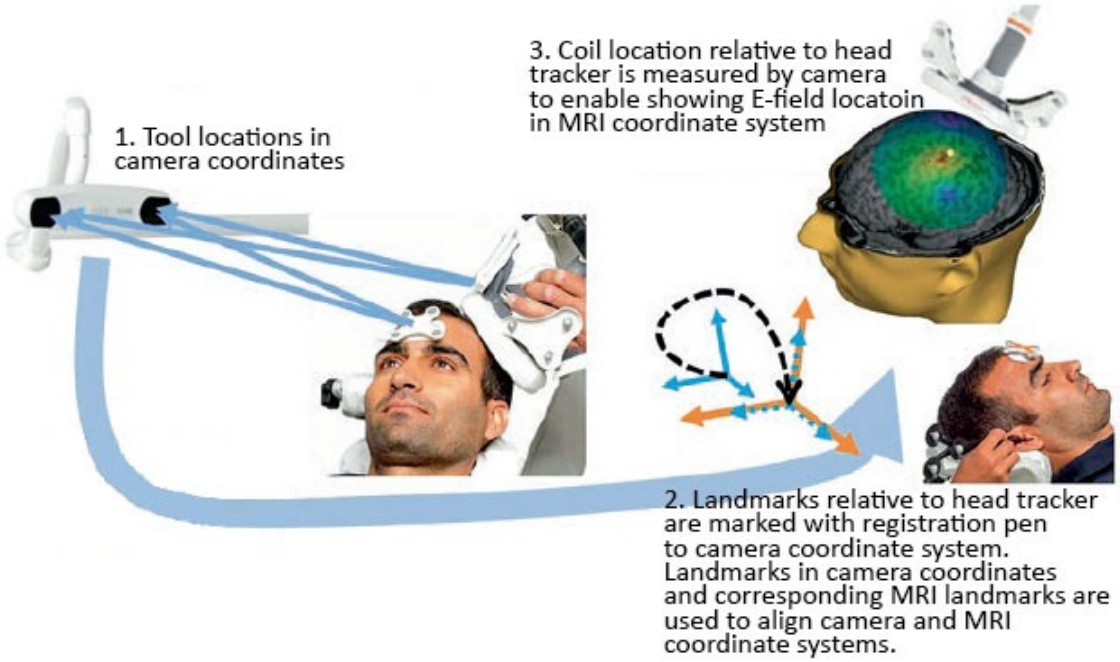


Figure 2: Electric field navigation [10]

2.3 Products

Nexstim Plc produces nTMS devices both for diagnostic and therapeutic applications. The diagnostic device called NBS is used for pre-surgical mapping of the speech and motor cortices of the brain. The brain is a plastic organ and functional areas can move into different locations when a tumor grows inside the brain [10]. Structural brain scans can not reveal the location of functional brain areas, and therefore nTMS is needed. Mapping the brain with nTMS gives knowledge of the locations of functional areas and thus makes surgery planning easier and safer. In mapping, hundreds of stimuli are given to the area close to the tumor or other areas of interest and electromyography (EMG) or speech tracking is used to find out, if stimulation excites a response in that area. The excitatory and inhibitory stimuli are marked to the MRI scans making a map representing safely operative brain areas and areas that should be preserved.

Nexstim's NBS product was Conformité Europeénne (CE) marked in 2003 to enable sales and marketing in Europe. Six years later the Food and Drug Administration (FDA) in the United States (US) approved the device. The latest approval for the diagnostic device was received from Health Canada in December 2019.

The therapeutic device, NBT, is used for the treatment of major depressive disorder and chronic neuropathic pain. About 30% of major depressive disorder patients fail to respond to standard medication treatment [12] and therefore alternative treatment methods are needed. Functional brain imaging has shown reduced activity in the dorsolateral prefrontal cortex (DLPFC) of patients suffering from depression [13] [14]. The nTMS depression treatment stimulates the DLPFC at 10 Hz or 50 Hz frequency

to increase its activity.

The medical treatment of neuropathic pain is far from optimal, as less than half of the patients get satisfactory pain relief from traditional treatment. It has been studied that high frequency rTMS stimulation of the primary motor cortex contralateral to the pain has analgesic effects. [15]

The intensity of the treatment is determined individually for each patient by finding out a threshold intensity at which a muscle responds to the stimulation. In pain therapy, the muscle is at the pain site and in depression therapy, it is typically a thumb muscle. The threshold intensity is determined by attaching EMG electrodes to the tracked muscle and giving pulses to the representative area in the brain motor cortex. A motor threshold algorithm detects responses and iterates the intensity to be the minimum value at which the muscle reacts to the stimulus. The stimulation intensity in treatment is determined as % of the motor threshold which makes the treatment individual for each patient.

In Europe, the NBT device was CE marked for the treatment of major depressive disorder and chronic neuropathic pain in 2012. The therapeutic device got FDA clearance for the treatment of depression in early 2018, enabling marketing and commercial distribution of NBT in the US. In 2019, NBT got a Health Canada license for treatment of depression.

Both of the company's devices consist of a TMS stimulator, figure-eight-shaped stimulation coil, coil holder, camera, EMG, workstation PC and display, foot pedal, patient chair, head tracker and digitize pen. In addition to those components, the diagnostic device contains an additional PC display and a speech mapping module, if speech license is purchased. Figure 3 presents the therapeutic device.

Figure 3: Nexstim's Navigated Brain Therapy (NBT) system.

2.4 New Product

The NBT system introduced earlier is Nexstim's first product targeted for therapy use. Its software is heavily based on the pre-surgical device software which is clearly seen in the graphical user interface (GUI). The GUI of NBT is presented in Figure 4. In NBT, everything can be done on one screen and the software does not give any hints on what the user should do next. As the depression market has grown, the typical user is not a researcher used to technical equipment but for example a psychiatrist. This has led to long training times and extensive usage of training resources.

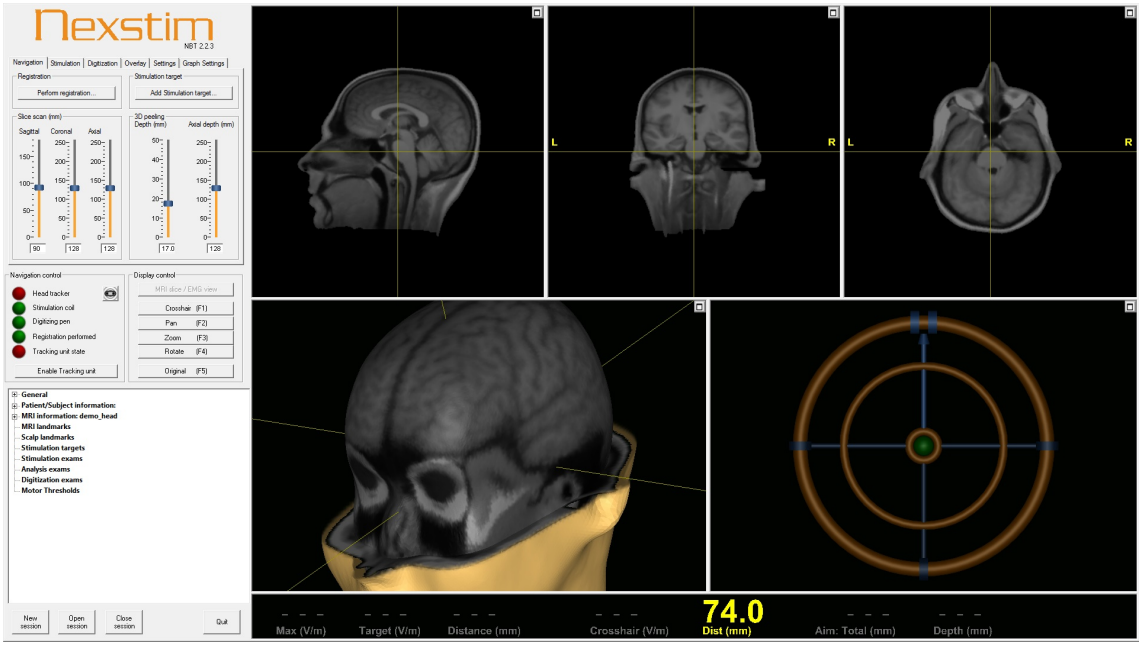


Figure 4: Graphical user interface of NBT software.

Now, the company's aim is to develop the next-generation of therapy product that is specifically intended for the treatment of major depressive disorder and chronic neuropathic pain. The new product will have a completely new, modern and workflow-driven GUI that can be operated from a touch screen. The depression workflow will be as straightforward as possible and designed only for the most basic use. The pain workflow will be slightly more open and suitable also for research purposes. In the first release, the GUI will be in English, but later it can possibly be localized to different languages. The aim is that the product will be easy-to-learn and easy-to-use. In this thesis, the new product will be called SNBT standing for smart navigated brain therapy.

In the development phase, the company specified usability goals for the new product. From the user's perspective, the system should have a guided process that minimizes mistakes and supports natural and easy workflow. Thus, the first usability goal was that the user must know what to do next. The user should be able to follow the right workflow path without mistakes or major error button pushes. The

second goal was that the user must be able to use the system from two different distances: by the workstation and by the patient. In several workflow steps the user is standing by the patient and needs to hear and see information without walking to the workstation.

For the customer, product simplicity and treatment effectiveness create resource efficiency. That is why the third usability goal was that treatment session should be carried out in less than one hour. The fourth goal was that the user shall be able to measure the quality of a treatment.

For Nexstim, the new product could amplify predictable sales and reduce the role of technical support in troubleshooting problems. The fifth goal was that user can explain verbally on what state the GUI is in and can troubleshoot on their own. The sixth goal was to decrease the time the system was not used by customers. That can be achieved, for example, by user knowing when their coils will expire and when to order yearly maintenance.

The new product will be CE marked and it shall meet the requirements that are needed for FDA's pre-market approval (510k). The approval processes are explained in Section 3.

3 Medical device regulation

This section presents the outline of medical device regulations in different market areas. Medical devices are regulated by national governments and international authorities to give the best possible treatment for patients. The control ensures that that no compromises are made in the safety, performance, or effectiveness of medical devices. The EU, US, and Canada regulations are in focus in this section as those are the target market areas of the new product. A data gathering table of the medical device regulation in the target areas is presented in Table 1.

3.1 European Union

Before 1990, all EU member countries had their own way of regulating medical devices [3]. In the 1990s, the Council of the European Union introduced three directives to harmonize the complex regulatory systems. The 1990 directive, AIMD 90/385/EEC, considered the legislation of active implantable medical devices, the 1993 directive, MDD 93/42/EEC, concerned medical devices, and the 1998 directive, IVDD 98/79/EC, in vitro diagnostic devices [4, 5, 6]. In 2017, the European Parliament and Council published Regulation 2017/745 on medical devices (MDR) and Regulation 2017/746 on in vitro diagnostic medical devices, repealing the previous three directives [16, 17]. The regulations do not have to be made into national laws unlike directives, and thus, the risk of different interpretations is reduced across the member countries. The new MDR has the same basic regulatory requirements as the old directives, but there are new requirements as well.

The MDR defines the "general requirements" to ensure the safety and effectiveness of a medical device whereas the old medical device directive required only safety and performance as described in the intended use of the device. The MDD and MDR both require manufacturers to conduct post-market surveillance (PMS), but MDR provides detailed requirements for the PMS. [5, 16] Considering usability, the general requirements of MDR state that the risks related to use error shall be eliminated or reduced as far as possible.

The MDR entered into force on May 26th, 2017 with a transition period of three years. From 2020, all new CE mark requests need to be delivered according to the new regulation, and CE marks delivered under the old directives can be valid until 2024. [16]

EU member states designate a governmental body called the Competent Authority to supervise the implementation of the MDR. The Competent Authority in Finland is Fimea. When a device fulfills the general requirements and the manufacturer has declared conformity with the MDR, a CE mark can be placed on the device. After earning the CE mark, the medical device can be marketed in all EU member countries. [16]

The MDR classifies medical devices into four risk-based classes (I, IIa, IIb, III) that have different demands to prove the conformance with the general requirements. Class-I devices are simple in design and pose extremely little risk of harm. [3] Class-I devices are further divided into sterile (Is), measuring (Im), reusable surgical (Ir) and

non-sterile, non-measuring or non-reusable surgical device classes [16]. Examples of different Class-I devices include scalpels for Ir, examination gloves for Is, stethoscopes for measuring and knee supports for non-sterile, non-measuring and non-reusable surgical. The manufacturers of all Class-I devices need to implement quality management system (QMS) and provide technical documentation that provide detailed information about the device. The QMS can be established by applying harmonized standard ISO 13485 *Medical devices – Quality management systems – Requirements for regulatory purposes* [18]. After the required documentation is done, the non-measuring, non-sterile or non-reusable surgical Class-I device producer prepares a declaration of conformity and is allowed to place a CE marking on the device. The QMS and technical file of the special Class-I devices need to be audited similarly as Class-II devices presented below. [16]

The conformity assessment of higher risk devices is overseen by Notified Bodies designated by the Competent Authorities. Notified Bodies are companies that specialize in evaluating products for CE marks. Class-IIa devices, such as endoscopes and powered wheelchairs pose a relatively low risk to the human body. Nexstim's NBS and NBT devices are in Class-IIa. Class-IIb products are medium to high-risk devices such as dialyzers, and orthopedic implants. [3] For the conformity assessment by a Notified Body, the company producing Class-II devices needs to provide technical documentation and implement the QMS [16]. If the assessment is successful, the Notified Body will issue a CE marking certificate for the device and a certificate of conformity for the manufacturer for complying with the quality requirements. The manufacturer of Class-II device needs to have the certificate before they can declare conformity with MDR.

Class-III products, such as aneurysm clips and artificial heart valves, pose a high risk to the human body [3]. For the conformity assessment, the Class-III device manufacturer needs to provide the same documentation as for Class-II devices. In addition, implantable Class-II devices and all Class-III devices require clinical investigation to prove the safety and performance requirements. The clinical investigation can be avoided if the device is equivalent to an already marketed device. In that case, the equivalence needs to be demonstrated to the Notified Body. [16]

3.2 United States

In the US, the safety of food and drugs became under regulation for the first time in the signing of the Pure Food and Drugs Act in 1906. This led to the establishment of the supervising authority FDA in the same year. The Act defined medical devices as drugs, but the requirements for drugs were not sufficient to assure the safety of devices. The definition of a medical device was added when the Act was modernized into the Federal Food, Drug, and Cosmetics Act in 1938. The 1938 Act entitled the FDA to bring charges against devices that were unsafe but no pre-market actions were required. [2]

With Medical Device Amendments Act (MDAA) approved in 1976, the FDA was finally justified to regulate the products prior to going to the US market. [19] The MDAA stated that manufacturers need to provide assurance that the medical devices

are both safe and effective [20]. The MDAA classified medical devices into three risk-based classes. Class-I devices present minimal risk of harm and are simple in design. [19, 21, 22] Class-I products include, for example, stethoscopes and surgical instruments. The regulatory pathway for Class-I devices is the easiest. The safety and effectiveness of Class-I devices can be ensured by so-called general controls. General controls include provisions of misbranding, adulteration, device banning, pre-market notification and so on [20]. However, most Class-I products are exempt from the pre-market notification [3, 22]. The FDA requires manufacturers of all medical devices to apply federal regulation 21 Quality System Regulation, but some Class-I devices are excluded from the process [23].

Class-II devices, such as computed tomography scanners, have a moderate risk of harm. Nexstim's NBS and NBT devices are in Class-II. For them, the general controls are not sufficient enough to prove the safety and effectiveness. Class-II devices are subject to special controls, which include performance standards, post-market surveillance, patient registries, special labeling requirements, pre-market data requirements, and guidelines. [24] Pre-market notification, better known as a 510(k) submission, is required for the Class-II devices. In the 510(k) submission, the manufacturer needs to demonstrate that the product is substantially equivalent to a previously cleared device. [3, 21, 22] Substantially equivalent devices have the same intended use, technical characteristics, and safety and effectiveness. Additional clinical data is usually not required for a 510(k) submission. Most new medical devices proceed to market through the 510(k) process. [3, 22]

Class-III devices have the highest risk of illness or injury and are therefore subject to the strictest regulation controls. Class-III devices include risky, completely new, and/or life-supporting devices, such as pacemakers and silicone breast implants. For Class-III devices, the general and special controls are not enough to prove that the device is safe and effective for its intended use. In addition to the controls, Class-III devices require pre-market approval (PMA) from the FDA. In PMA, the manufacturer demonstrates the safety and effectiveness of the device through clinical studies. [3, 21, 22, 25] All devices without a predicate are classified as Class-III devices even though the risk might be low. The manufacturer of the device can request for reclassification into Class-I or Class-II through the De Novo process. [3, 26] There are two requirements for devices seeking for the De Novo requests. The device can not belong to a device type that has previously been classified and the general or special controls need to be sufficient to prove the safety and effectiveness of the device [22]. De Novo devices are exempt from the PMA process, but 510(k) might be required.

3.3 Canada

In Canada, medical device quality, safety, and effectiveness requirements are defined by the Food and Drug Act and medical device regulations [27]. The medical device regulation authority in Canada is Health Canada working under the health ministry. Inside Health Canada are several branches and operational directorates of which Therapeutic Products Directorate (TPD) is responsible for medical devices. The TPD

supervises that medical devices entering the Canadian market fulfill the requirements of the Food and Drug Act and regulations. [28]

The classification of medical devices is similar but not identical to the classification system in the EU. The classification is risk-based and has four classes: I, II, III, IV. [28] Before marketing in Canada, the manufacturer of Class-I device needs to apply for an Establishment Licence. Manufacturers of Class-II, III, IV devices need a Medical Device Licence before they can market and sell their device in Canada. As in the approval processes in other countries, the number of requirements varies between device classes. [29] One requirement for the Class-II, III and IV device manufacturers is to implement QMS according to ISO 13485 and additional requirements from Canadian Medical Device Regulation. The requirements for Class-II devices are equivalent to the 510(k) in the US and Class-IV requirements equivalent to PMA.

In Canada, Nexstim's NBS product is classified into Class-II as it is an active diagnostic device intended to image physiological features. Active therapeutic devices intended to administer energy to the patient's body are typically classified into Class-II, but as the company's NBT device could be potentially hazardous, due to the repetitive TMS pulses, it is labeled into Class-III.

3.4 Other areas

In 2010, the World Health Organization (WHO) reported that approximately 30% of countries have a framework for regulating medical devices, 30% have partial regulation and the rest are developing a system or do not have any regulation [30]. As the global interest in medical devices has increased and many countries still lack regulatory frameworks, WHO developed a model of a global regulatory framework for medical devices and in vitro diagnostic devices. [31] The publication aims to guide countries in develop a regulatory system.

As in the EU, the medical device regulation is the same in all member countries, most of the non-EU countries have their own regulations and approval procedures that have been developed before WHO's guidelines. Some countries have approved international regulations and standards, but they still might have additional laws and rules that are not in line with international regulations.

For example, in Australia, Therapeutic Goods (Medical Devices) Amendment Regulation was originally based on the EU directives on medical devices, but as it has been further developed, it now contains several differences. The risk-based device classification is similar to the EU one, but it has five classes. Class-I for non-sterile and non-measuring devices, Class-II for sterile and measuring, and Classes IIa, IIb, and III. The quality management system required for other devices than Class-I is based on the international standard ISO 13485. [28]

Some countries, such as Brazil, require manufacturers to have a local distributor or representative that is responsible for issuing applications and technical documents to the authorities. In Brazil, the device classes are I, II, III, and IV which are similar to those in the EU. Devices in classes I and II have two possible regulatory pathways. "Cadastro" is the faster and easier procedure for low risk devices that have been separately listed in Brazilian medical device law. "Registro" is a more demanding

pathway for high risk devices. Class-III and IV devices have to go through the Registro pathway and an audit to get a Brazilian Good Manufacturing Practices certificate. In Registro procedure, manufacturers need to provide full technical documentation and in some cases also clinical data. All documents have to be written in Brazilian Portuguese. [28]

In China, the medical device regulation field is quite complex and different provinces might have different practices. The classification system is similar to the EU, but it only has three classes. A device manufacturer needs to provide a quality certificate according to the international quality management standard, ISO 13485. Before approval, Class-II and III devices need to be tested in Chinese test laboratories and clinical studies performed in China might also be required. All documentation for the application needs to be provided in simplified Chinese. [28]

Region	Legislation	Implementation of QMS	Device Classes	Regulatory Authority
US	Medical Device Amendments Act (MDAA)	FDA Quality system regulation	I, II, III	FDA
EU	Medical Device Regulation	ISO 13485 standard	I, Is, Im, Ir, IIa, IIb, III	Notified Body designated by national Competent Authority
Canada	Food and Drug Act	ISO 13485 + specific requirements	I, II, III, IV	TPD under Health Canada

Table 1: Summary of medical device regulation in the US, EU and Canada.

4 Usability standards

The regulation of medical devices has improved over the years but they have not been safe enough and use errors of such devices became a frequent concern. In 1993, the first national usability standard for medical devices was published by the Association for the Advancement of Medical Instrumentation (AAMI) in the USA. The standard, AAMI HE48:1993 *Human Factors Engineering Guidelines and Preferred Practices for the Design of Medical Devices*, adopted the usability guidelines of an existing standard, MIL-STD-1472 – *Human Factors Engineering*, written for military equipment. The medical device usability standard covered introduction and guidelines to human factors design. [32]

In 2001, AAMI published a new standard partially replacing the 1993 guidelines. The AAMI HE 74:2001 *Human factors design process for medical devices*, focused on the user-centered process that was linked to the requirements set by the FDA. This standard was later adopted by the International Electrotechnical Commission and released as IEC 60601-1-6:2004 *Medical electrical equipment – Part 1-6: General requirements for safety – Collateral standard: Usability.* [32] Since then, the international standard has been updated three times and it has been separated into two different standards: the IEC 60601-1-6 and IEC 62366 *Application of usability engineering process to medical devices* [33].

IEC 62366 was originally published in 2007 to separate the usability engineering process from the second edition of IEC 60601-1-6. In 2015, the second version of IEC 62366 was issued. Both IEC 60601-1-6:2013 and IEC 62366:2015 standards are valid today and approved in the EU, Canada and the US [34, 35, 36].

The IEC 60601-1-6:2013 describes the general requirements for medical device safety. The standard states that "The medical electrical equipment shall provide adequate usability such that the risks resulting from normal use and use error are acceptable". [33] Here the medical electrical equipment denotes a medical device. The second requirement of IEC 60601-1-6:2013 is that a usability engineering process described in international standard IEC 62366-1:2007 shall be performed [33]. Since then, the second version of the IEC 62366 standard has been published, but 60601-1-6 has not been updated to take that into account. The usability engineering process described in this thesis is applied by the latest 2015 version of the IEC 62366.

4.1 IEC 62366-1:2015 Application of usability engineering to medical devices

The IEC 62366 standard has two parts. Part 1, published in 2015, describes the usability engineering process and part 2, published in 2016, gives guidance on how to conduct the process. The usability engineering process is intended to identify and minimize the use errors and thereby reduce use-associated risks. The standard describes how a manufacturer should analyze, develop, and evaluate the safety-related usability of the medical device. The general requirements of the standard state that the process shall address not only the user interactions with the software but also, the interactions related to device transport, storage, installation, operation, maintenance,

repair, and disposal. As software is the new component of the new SNBT system, the evaluation part of this master's thesis focuses on software.

The standard requires the manufacturer to reduce the use-related risks by adapting safety by design, implementing protective measures in the medical device itself or in the manufacturing process or creating information for safety. Information for safety includes a user manual, warnings, and labels on the device. The usability engineering process needs to be applied also to the information for safety to determine that the material is perceivable, understandable and supports correct use. The extent of the process is determined by the complexity of the user interface (UI) and the severity of the harm associated with the device.

4.1.1 Requirements of the usability engineering process

The usability engineering (UE) process lists multiple steps the manufacturer needs to go through to prove that the medical device is easy to use. The steps of the UE process are presented in Figure 5. The process requires the manufacturer to create a usability engineering file that contains all documents delivered.

Clause 5.1 of IEC 62366:2015 states that the manufacturer must prepare a use specification for the medical device. The use specification must include the intended use of the medical device and definitions for a patient, user, and use environment. The document needs to describe which part of the body or tissue type is interacted with as well as the operating principle of the device.

Clause 5.2 of the standard requires the manufacturer to identify UI characteristics and potential use errors related to safety. The characteristics shall be determined by the application of risk analysis described in the risk management standard ISO 14971:2007. As part of the analysis, the producer shall identify known and foreseeable hazards and hazardous situations that could affect the patient's or user's safety (Clause 5.3 in IEC 62366:2015). Clause 5.4 requires that the hazard-related use scenarios and the severity of the associated harm need to be described. After all hazards and risks arising from them have been analyzed, the manufacturer shall select the hazard-related use scenarios for summative evaluation, known previously as usability validation (Clause 5.5).

Technical requirements for the UI need to be specified according to Clause 5.6 of IEC 62366:2015. These requirements can include specification, for example, for the character size on the screen and sound volume used in the UI. At this phase the manufacturer needs to decide whether accompanying documentation and medical-device-specific training is required.

The usability of the device needs to be tested through the development process and after the product is ready to be marketed. The assessment done during the development is called formative evaluation and the assessment done for the final product is summative evaluation. Clause 5.7 requires the evaluation plans to document the objective of testing, used methods and which part of the UI is being evaluated. If usability tests are done, the plans should include a description of the test environment and participants. It also needs to be determined whether the participants need training before the test and is there any accompanying documentation

given to them during the evaluation. In addition to these, the formative evaluation plan should document when in the UE process the assessment should be performed. The summative evaluation plan shall specify criteria for determining whether the information for safety is perceivable, understandable and supports correct use.

After the evaluation is planned, the manufacturer needs to design and implement the UI (Clause 5.8). At this point, the accompanying documentation and training material should be established. The formative evaluation is performed at different implementation phases and the results are stored in the UE file. If use errors, hazards or hazardous situations are discovered during the assessment, steps of the UE process shall be repeated.

The final step of the UE process is to perform the summative evaluation for the final UI (Clause 5.9). Data from the assessment shall be analyzed to identify the consequences of all use errors emerged. If new use errors, hazards or hazardous situations are discovered during the analysis, the manufacturer must re-enter the UE process. If no new hazards are found, the company needs to determine whether further improvements are necessary. If improvements are not necessary, it needs to be reasoned. The final step is to determine the residual risk related to use and evaluate it according to the risk management process defined by ISO 14971:2007.

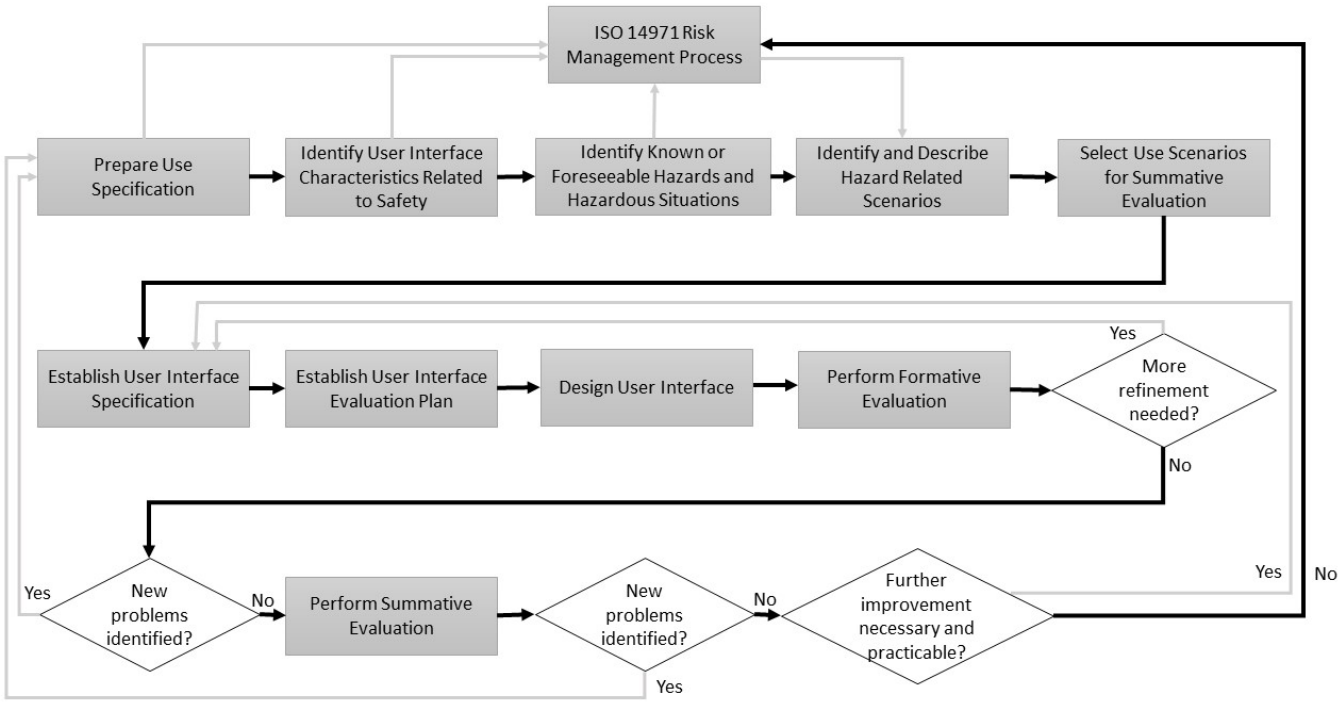


Figure 5: Usability Engineering process determined in IEC 62366-1:2015. Some parts of the usability engineering process give input to the risk management process described in standard ISO 14971. [8]

5 The quality system of the company

This Section presents the quality management system of the company for which the UE process is applied and describes how the system was updated to comply with the new MDR.

The fundamentals and guidelines for Nexstim's quality management system and business are defined by the company's quality policy and values. The policy aims to assure that products and services meet the customers' needs, are competitive, are delivered to the customers as promised, and adhere to legal requirements and applicable standards. The company's internal values include economic efficiency, innovativeness and personnel well-being. Customer values include efficiency and reliability, confidence in products and processes, and customer satisfaction.

The quality policy and objectives are achieved by following the company's quality manual approved by the Head of quality assurance and regulatory affairs (QA&RA). Nexstim's QMS adheres to the requirements of the international ISO 13485:2016 Quality Management System standard, FDA Quality System Regulation, the European Medical Device Directives and Regulation, Canadian Medical Device Regulations, and Australian Therapeutic Goods Regulations. The European directives are complied with until the firm's QMS is refined to follow the new MDR in 2024 at the latest. The quality system applies to the design, manufacturing, marketing, sales, after-sales, and service of Nexstim's medical devices.

The quality system of the company is divided into three document groups: quality system documents, product-specific documents, and general quality system records. The quality system documents create the backbone for quality by defining the quality policy and manual and the operating instructions for performing different functions for proofing excellence. Operating instructions, also known as standard operating procedures (SOPs), describe the actions and documents the manufacturer needs to provide to comply with, for example, a specific standard or a regulation. The company's SOPs include instructions, for example, for conducting a clinical evaluation and performing a risk management process.

The application of the quality system produces multiple product-specific documents that need to be stored. The design history file (DHF) stores all documents that are created through the development process of a product. DHF documents include, for example, project plans, different level requirements for the product, and test plans. The device master record (DMR) consists of the documentation needed to manufacture the device. The DMR includes instructions, for example, for assembling and packaging. Finally, the device history record (DHR) stores the documentation generated during the manufacturing of an individual device. DHR demonstrates that the device was assembled in accordance with the DMR.

The general quality system records are plans and reports of achieved results or certifications of performed actions. The records include for example internal audit reports, customer complaints, and training requirements for employees.

5.1 Updating the quality system

In the earlier Section, it was presented that considering usability, the general requirements of the MDR state that the risks related to use error should be eliminated or reduced as far as possible. The reduction of use errors can be shown by applying the international harmonized usability standard IEC 62366 presented in Section 4.

In the company's quality system, instructions for performing the UE process described by the usability standard have been presented in an SOP Q310 Usability Engineering Process established according to the first version of the standard. The SOP has been outdated since the new version was published in 2015.

This thesis developed a new version of the SOP to comply with the newest version of the standard. The new version of the UE process SOP lists the requirements of the usability standard and the deliverable documents to prove accordance with each requirement.

The documents resulting from the usability engineering process are stored in a usability engineering file which is part of the products DHF except for user manual and training material being part of the DMR.

6 Results of the usability engineering process

This section describes the application of the usability engineering process for the new product. The company manufacturing the device has performed the process before for their existing devices, but it has been done according to the previous versions of the IEC 62366 usability standard. In this thesis, the process was updated to meet the requirements of the newest version of the standard.

6.1 Use specification

The use specification was created in collaboration with the company's application specialists, physician, and project assistant. The generated use specification is presented in the following subsections.

6.1.1 Intended medical indication

The device will be used to treat major depressive disorder and chronic neuropathic pain, thus the intended medical indication was specified separately for the two applications. For the depression application, the system is intended to be used for the treatment of major depressive disorder by targeting and delivering non-invasive, repetitive, nTMS stimulation to the patient's brain region called the dorsolateral prefrontal cortex. The system is indicated to treat adult patients who have failed to receive satisfactory improvement from prior antidepressant medication in the current episode.

For the pain application, the system is intended to be used for the treatment of chronic neuropathic pain by delivering non-invasive, repetitive, nTMS stimulation to the patient's motor cortex. The pain application is indicated to treat chronic neuropathic pain in adult patients to alleviate pain.

6.1.2 Intended patient population

The intended patient population was decided to be adult patients who weigh less than 200kg. The weight limitation came from the technical specification of the patient chair. For safety reasons, the device is not intended to be used with persons with non-removable conductive, ferromagnetic or other magnetic-sensitive metal anywhere in the head or within 30 cm of the stimulation coil. Using TMS with patients with metal in the head is prohibited due to the attractive and repulsive forces generated by the magnetic field that could move the metals [37].

Patients who have an active or inactive implanted device including deep brain stimulators, cochlear implants, cardiac pacemakers, and vagus nerve stimulators are not in the intended patient population because of possible heating and stimulation of the devices. The devices could heat up due to the magnetic field induced eddy currents in the conductive surface electrodes and implants. Heating brain tissue above 43 °C can result in permanent damage. Furthermore, the implanted devices could fire unintended stimulation, if voltages are induced to the electrode wires. [37]

The most serious adverse effect of rTMS is seizures, although the risk is minimal [37]. Seizures are more dangerous to some people and therefore the device should not be used with patients with increased intracranial pressure or serious heart disease [38]. Several factors can also increase the risk of seizure. These factors include, for example, substance abuse and intake or withdrawal of some drugs affecting the nervous system [37]. That is part of the reason why TMS should not be used without a clear benefit for patients suffering from other mental illnesses than depression.

The efficacy of rTMS therapy has not been proven with certain patient populations. These include people suffering from seasonal affective disorder, substance-induced depression, and depression secondary to a general medical condition. Several TMS studies have excluded subjects that have not received satisfactory improvement from antidepressant medication, have a recent suicide attempt or have a history of unsatisfactory treatment with electroconvulsive therapy or vagus nerve stimulation. The studies have included subject population aged 22 to 70 years and pregnant women have been excluded. The depression treatment effectiveness can not be stated for populations that have not been studied.

6.1.3 Intended part of the body and tissue interacted with

The TMS coil is placed on the patient's scalp and the stimulation affects the tissues in the scalp and cortical surface. The device contains an EMG system which electrodes are placed on skin over muscle.

6.1.4 Intended user profile and use environment

The potential users of the device include psychiatrists, neurologists, nurses, hospitals' biomedical engineers, and the technical support personnel of the manufacturer. The users can be grouped as Users, including clinical personnel, Admin Users, including hospitals' biomedical engineer, and Service Users such as technical support personnel. The users of the system shall interact with the device depending on their specific roles. Some will be responsible for operating the device on patients, others will execute administrative and service tasks on the system software and hardware.

All users shall have at least the following general skills: understand English, read and understand the Latin alphabet and the westernized Arabic numerals and have basic computer skills such as operating a touch screen, keyboard and mouse and using external storage devices. The user is permitted to have mild vision and hearing impairments, but the device is not intended to be operated by persons who are pregnant, have any electrical implants or have conductive, ferromagnetic, or other magnetic-sensitive metal implants anywhere in the upper body.

The level of education may vary between users, but only qualified medical personnel should use the device on patients. All users should go through at least the initial device training.

The device is intended to be used in hospitals or other health facilities. The system is movable within the same room and building.

6.1.5 Operating principle

The system combines non-invasive TMS with MRI-based stereotactic navigation and simultaneous surface EMG. The system uses stereotactic localization of the stimulation coil to interactively visualize the calculated effects of TMS stimulation and to guide TMS precisely over cortical areas.

6.2 User interface characteristics and potential use errors related to safety

This section presents an analysis of the user interface characteristics and potential use errors that could affect the patient's or the user's safety. The investigation was done using two different techniques: risk screening and task analysis. Identifying UI characteristics by going through a questionnaire was required by the usability standard, but there were several options for analyzing possible use errors. Task analysis was selected as it was the most analytic form of investigation suggested by part two of the standard. The aim of the investigation was to find factors that could be potential sources of injury or damage and that should be further evaluated in the risk analysis presented in Section 6.3. In risk analysis, physical injuries and damage to health or property are called harms and potential sources of harm are called hazards.

6.2.1 Risk screening

The safety-related UI characteristics of the new product were screened by going through the UI related questions in the risk management questionnaire described in ISO 14971:2007 Annex C [39]. The purpose of the screening was to identify the characteristics that could cause possible hazards and that should be further evaluated in the risk analysis. The safety questionnaire contains a total of 34 questions of which nine are UI related. The UI questions try to identify whether the successful application of the device depends critically on human factors.

The successful use of the new SNBT device depends on the GUI design. Users need to interpret displays correctly in order to treat patients and avoid use errors. Weak GUI design can be a source of use errors.

Distractions during the use of a device can lead to use errors. However, the company's new product is meant to be used in a TMS room where disturbances are untypical and thus the possibility of hazards arising from interruptions is minimal.

Connecting accessories to the SNBT device can create multiple possible hazards if items are not connected properly. For example, attaching a damaged coil to the stimulator can cause sparking and lead to the user being exposed to high temperatures creating a risk of burn. However, connecting accessories is not related to the system software and hence it is not further analyzed here.

The successful application of the new device includes using control interfaces as the application must be operated via a touchscreen, mouse, keyboard, and foot pedal. From these, the foot pedal is the most probable source of harm as it is used to initiate stimulation.

The SNBT device GUI displays information that needs to be interpreted correctly to enable proper use of the system. The safety-related UI characteristics include displaying electric field distribution, EMG responses, and coil positioning during treatment. Misunderstanding data is a use error that could possibly lead to the safety of the patient being compromised.

The new product has simple menus that will prevent the user from seeing the underlying screen. In the main menu, the user can read the screen related contextual help, take screenshots, adjust the sound volume et cetera. The possible hazards arising from the menu design should be further analyzed.

The users' background and skills play an important role when considering a successful application of the new product. The intended users of the system are described in the use specification presented in Section 6.1 and no potential sources of harm should be created if the specification is followed.

The UI of the new system is used to initiate user actions and hence it is important to display explanatory information to avoid use errors. One example of a UI initiated action is that the GUI shows if the stimulation is not hitting the target. Then, the user has to act and adjust the coil location.

Devices that depend on the essential performance of a function can cause great risk of harm to the user's or the patient's safety if a fault arises. These kinds of functions are, for example, alarms and life supportive devices display information. The new device considered in this thesis does not depend on the essential performance of functions and thus, it is not taken as part of the risk analysis.

The risk screening resulted in five GUI related points that should be taken into account in the risk analysis. The GUI characteristics for the further analysis include design features contributing to use errors, control interfaces, displaying information, menu, and initiation of user actions from GUI.

6.2.2 Task analysis

Potential use errors that could affect safety were identified by conducting a task analysis. Safety-related use errors are hazards that should be further analyzed in the risk analysis. Task analysis is done by going through all the user interactions with the medical device to develop an understanding of factors that could hinder user performance. All functions of the device were vetted with the help of workflow charts that had been formed earlier in the GUI development process. Using the system involves several tasks and thus the task analysis was done separately for different task levels.

Use cases presented in Figure 6 are the highest-level tasks in the system. It was identified that from the use cases, only the following could lead to potential use errors related to safety: perform examination and therapy in depression application, perform examination and therapy in pain application, move the system to another room, clean patient contact surfaces, and move the patient into and out of the chair. Furthermore, when analyzing the lower-level tasks in these use cases, the use errors arising from the software were identified to appear when performing examination and therapy in depression and pain applications. The pain examination and treatment include

almost the same set of sub-tasks and possible use errors as depression application and therefore only the task analysis for depression application is presented below.

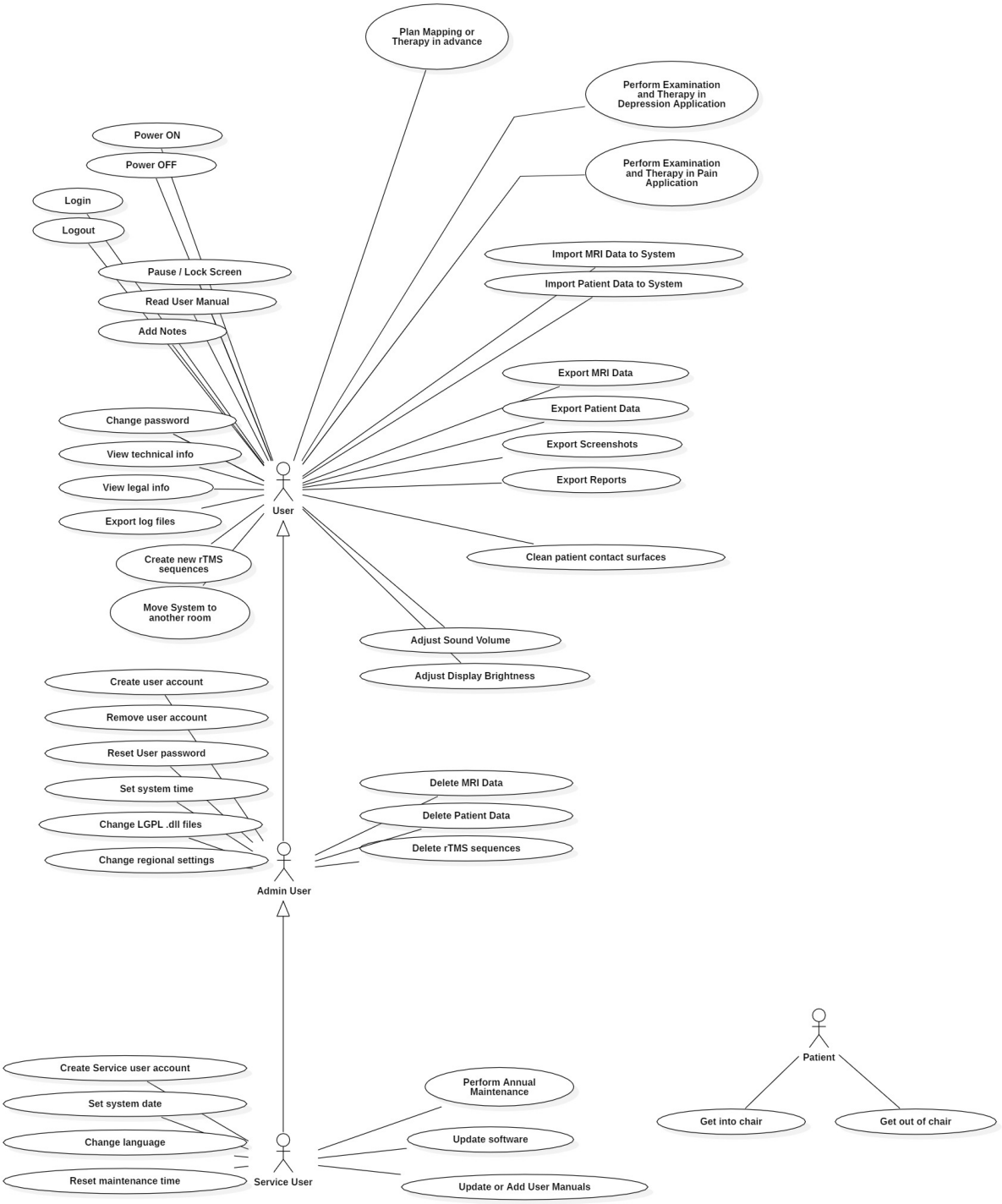


Figure 6: The use cases when using the system.

Perform examination and therapy in depression application: Perform examination and therapy for depression use case includes several sub-level tasks that are presented in Figure 7. After analyzing, it was discovered that all other sub-level tasks except for "Activate Head Tracker" could create software-related use errors affecting safety.

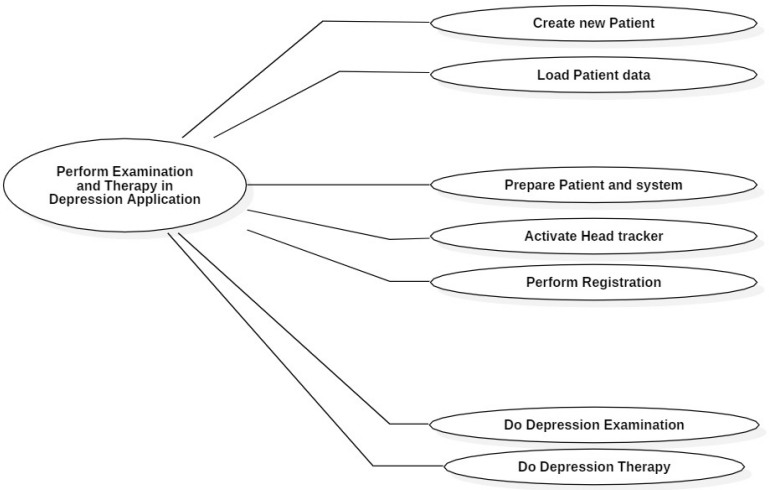


Figure 7: The sub-level tasks in depression workflow.

In 'create new patient' task, the user needs to select MRI for the patient, check that the 3D reconstruction of the MR images is accurate, enter patient information and set landmarks to the MR images. One possible use error is that the user could select a wrong MRI for the new patient and hence the head 3D model would not correspond with the actual head of the patient. Another safety-related use error is that the user could select incorrect places for the MRI landmarks which could lead to incorrect registration.

In 'load patient' task, the user needs to select a patient, treatment plan and whether the user wants to map the patient or give treatment. In addition, the user can also create new treatment plans and review the maps and given treatments. Use errors that could be made here include selecting the wrong patient, treatment plan or active map. Selecting the wrong map as active could lead to delivering treatment with too high or low intensity.

'Prepare patient' includes sitting the patient into the treatment chair, attaching EMG electrodes, and reviewing the EMG signal quality. Possible harm can be done to the patient if the user selects the wrong EMG channel for the motor threshold determination as then the treatment could be given with too low or high intensity. Another possible use error is that the user mislabels the EMG channel so that it does not correspond to the muscle the electrode is attached to. This could lead to an incorrect map of the representation area of the muscle.

'Perform registration' includes the following III-level tasks: set or reset MRI landmarks, basic registration, and advanced registration. Set or reset MRI landmarks involve the same possible use errors as the same task when creating patient. In basic

registration, the user needs to point the MRI landmark locations in the patient's head with a registration pen. However, if the user points wrong locations, the registration can be incorrect. In advanced registration, nine additional points are set to patient's head surface and careless work can similarly lead to incorrect registration.

In depression examination, the system operator needs to perform several tasks including mapping the motor cortex and selecting a pulse from the map, determining motor threshold (MT), and setting a treatment target. In all tasks involving stimulation, a possible use error is to hold the coil poorly so that the bottom of the coil is not tangentially positioned on the patient's head. Holding the coil poorly could lead to a too large area being stimulated and too high or low stimulation intensity due to an incorrect MT result. Another use error is creating a wrong target for the MT determination or for the treatment. Furthermore, the MT result could end up being incorrect, if the determination process is prematurely terminated.

When giving treatment to the patient, the user needs to start and monitor the treatment and view the treatment result. Optional tasks include re-determining MT, pausing and stopping treatment and selecting and creating new treatment targets. In addition to aforementioned use errors, giving treatment could result in the user selecting wrong repetitive TMS (rTMS) sequence for the therapy and treatment target not being reproduced. Insufficient treatment could be given if the stimulation coil moves out of target or the treatment is stopped prematurely. Selecting a wrong rTMS sequence can lead to stimulation with the wrong intensity and frequency.

Task analysis resulted in several possible use errors that could affect the patient's safety. These use errors are further analyzed in the next section.

6.3 Risk analysis

This section presents the risk analysis performed as required in the usability standard. The standard obligates the manufacturer to identify the foreseeable hazards, hazardous situations and use scenarios leading to them that can arise from the UI of the new product. Moreover, the severity of the harm that can emerge from a hazard-related use scenario needs to be assessed.

The risk management standard ISO 14971 defines a hazardous situation as a circumstance in which people, property or environment are exposed to one or more hazards [39]. The severity of a harm is assessed using a five-point scale where 5 denotes catastrophic harm that results in death and 1 denotes negligible harm that causes temporary discomfort or inconvenience. The severity scale is presented in Table 2.

The risk analysis of the new SNBT product was conducted using the analysis of the existing NBT device as a base. NBT uses almost the same hardware and most of the software functions are the same although the graphical representation is not identical. Thus, the hazards and hazardous situations arising from the two products are similar. Some of the possible hazards and hazard-related use scenarios of the new device were identified during the risk screening and task analysis presented in Sections 6.2.1 and 6.2.2.

As described earlier, hazards are potential sources of harm and injury. Hazard

Index value	Severity	Description
5	Catastrophic	Results in death(s)
4	Critical	Results in permanent impairment or life-threatening situation, if medical intervention is not obtained
3	Serious	Results in injury or impairment requiring professional medical intervention
2	Marginal	Results in temporary injury not requiring professional medical intervention
1	Negligible	Temporary discomfort or inconvenience

Table 2: Severity scale used in risk analysis

can not lead to harm on their own unless a sequence of events creates a hazardous situation. Hazard can be, for example, high temperature, but it does not cause a skin burn unless a person sips hot tea and is exposed to the high temperature. In this case, the use scenario is that a person sips hot tea, the situation is exposure to the hot temperature, and the harm is a skin burn. Skin burn is a temporary injury that does not require medical assistance and is hence ranked as severity level two.

In risk screening, it was evaluated that displaying important information is a factor that should be further examined in the risk analysis. One example of a hazard that can arise from displaying information is an attentional failure. If the use scenario is that the EMG responses and coil position is not visible to the user when standing next to the patient, the hazardous situation is that the system can not be used. In this case, the harm is that the system is unusable, and patients can not be treated. However, the severity ranking of this harm is only one as it will create only temporary discomfort for the patient.

The task analysis discovered possible use errors that could create greater and more severe harm. One example is in depression examination, where the user could end the MT determination prematurely which could lead to too high stimulation intensity due to too high MT. The hazard is that the patient is exposed to a magnetic field and the hazardous situation is that the safe amount of stimulation is exceeded. The harm that could arise from this situation is an epileptic seizure. Seizures are ranked with severity level three as they are injuries that require professional medical intervention.

The risk screening and task analysis revealed also completely new risks that had not been discovered during existing systems risk analysis. In the new software, the user must select a treatment plan from which she is giving treatment for the patient. An attentional failure is observed if the user selects a wrong treatment plan. The treatment plan contains the therapy targets and thus using the wrong plan can lead to stimulation of unintended areas during the treatment. The hazardous situation arising from this use scenario could be that the effect of the treatment can be decreased as stimulation is not given to the right location. The possible harm is ineffective treatment with a severity ranking of one, as it will only create inconvenience for the patient.

Another use scenario that had not been reported in the existing device's risk analysis is the user selecting the wrong rTMS sequence for the treatment. Incorrectly selected rTMS sequence results in too high or low stimulation intensity and frequency. If the parameters are too high, the hazardous situation is that the safe amount of magnetic stimulation is exceeded, and if they are too low, the effectiveness of the treatment can be decreased.

In total, the risk analysis resulted in over 120 risks related to usability and from those around 50 were software-related. Some hazard-related use scenarios were presented twice in the risk analysis as the same use scenario can lead to the amount of magnetic stimulation being too high or low. Examples of analyzed risks are presented in Table 3.

Hazard	Use scenario	Hazardous situation	Harm	Severity
Attentional failure	Needed information such as EMG responses and coil position during rTMS therapy are not visible to user	User can not use the system	System unusable	1
Magnetic field	User ends MT determination prematurely	Safe amount of magnetic stimulation exceeded in treatment due to too high MT value	Seizure or equivalent	3
Attentional failure	User accidentally selects the wrong treatment plan causing stimulation of unintended areas	Effect decreases due to stimulation of the wrong location	Ineffective treatment	1
Attentional failure	User selects the wrong rTMS sequence for treatment	Effect decreases due to stimulation with too low frequency	Ineffective treatment	1
Attentional failure	User opens the wrong patient	Safe amount of magnetic stimulation exceeded, stimulation of unintended areas	Seizure or equivalent	4

Table 3: Examples of risks evaluated in the risk analysis.

6.4 User interface specification

The usability standard requires the manufacturer to create user interface specification that includes testable technical requirements relevant to the UI and an indication

whether accompanying documentation and medical device-specific training is required. Accompanying documentation is considered as part of the product and that is why the UI requirements should include the specification for the user manual as well. In addition, the content of the training material should also be defined.

The UI specification was written within the system requirements. The system requirements include specifications for different hardware devices as well as the software in general. The system's usability requirements were written to ensure that the usability goals presented in Section 2.4 would be met. Each goal generated one to three requirements. The first goal was that the user must know what to do next. This generated requirements for the workflow-driven GUI and screen headings that describe what the user should do in the step.

The second goal, that aroused three requirements, was that the user must be able to use the system from two different distances. The first requirement was that the software shall give audio feedback during operations where it is likely that the user does not look at the screen. The second requirement was that the software GUI shall have big enough font size for information that the user needs to see when standing next to the patient. The third requirement was that the software shall support using the foot pedal in operations where the user is not at the workstation.

The third goal was that the treatment session should be carried out in less than one hour. The longest approved depression treatment lasts for 37 minutes and thus it was specified that opening data and preparing the patient shall be performed in less than 20 minutes.

The ability to measure treatment quality was the fourth goal which generated the requirement for the software GUI to present quality data after the treatment is given. The requirement developed from the fifth usability goal was that each screen shall have a unique heading to enable the user to know where one is in the workflow process. The final goal considered decreasing system downtime and created requirements for the system to inform the user about aging coils and approaching yearly maintenance in advance.

The accompanying documentation of the new product includes the user manual and a separate manual for the serviceman. The user manual of the existing NBT system was used as guidance when writing the specification for the new manual. The new product will have a traditional printed manual, as well as a software integrated manual and a workflow step specific contextual help that can be read from the workstation PC. The manual requirements list information that needs to be presented. The content is specified to include information about the safety aspects, used technologies and features, hardware devices and use instructions. For example, the user manual shall include information about the intended use and contraindications for using the system.

Like the user manual specification, the requirements for the training material specify what should be trained to the user. The training material has been divided into system training and application training. In system training, users are trained to take care of the system's daily maintenance and how to move the device as well as about contraindications for use. Application training covers specific tasks that users need to perform to treat patients.

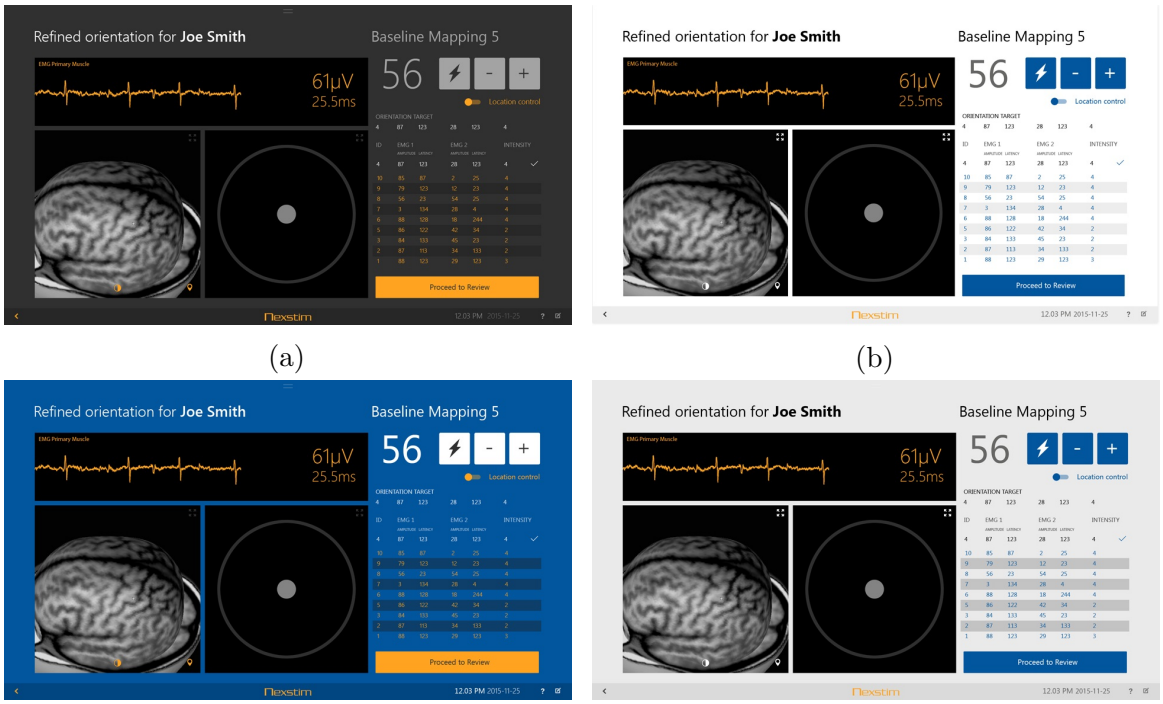


Figure 8: Color alternatives for SNBT GUI. The dark grey design presented in (a) was chosen for the product.

6.5 User interface design and implementation

The hardware of the new device was decided to be kept the same as the existing therapy device NBT and the GUI was the only component that needed to be redesigned. The GUI design process started in 2012 when an external company was hired to design the outlook of the software. Before the GUI could be designed, all functions and tasks related to the system usage needed to be analyzed. The tasks were linked into a chronological order in the treatment process to make an outline for the workflow-driven GUI. The screen design started from rough sketch models that had one main task on each screen. The design and screen contents were further developed aiming at a modern and easy-to-use GUI.

The design process faced difficulties and the company designing the GUI was changed two times. In 2015, a software design company was hired to redesign and complete the GUI. When the treatment workflow was finalized and screen components were decided, the GUI outlook and color were chosen. Figure 8 presents four color alternatives from which the darkest design, 8a was chosen.

The GUI design was an iterative process where workflow step screens were drawn, evaluated and improved according to feedback. Regular review meetings were attended by the project manager, project assistant, product manager, and graphical designers. When the screen design was suitable for the workflow step, the screen was accepted and given to software developers to implement.

Nexstim's existing software had been implemented using the C++ programming

language and Microsoft Foundation Class Library (MFC). C++ had been selected for its excellent performance in real-time calculations required, for example, for E-field visualization during stimulation. MFC is a C++ library that was used for developing the Windows-based GUI for the company's previous products.

MFC was discovered to be too stiff for developing the new workflow-driven GUI, and thus Qt and Qt Modeling Language (QML) were chosen for the GUI development. The choice of Qt and QML was reinforced by the fact that they provide a powerful modern software development framework and support using C++ natively. QML was chosen alongside Qt as it seemed at certain point that Qt would phase out the native desktop GUI components and QML was the upcoming primary language for flexible GUI implementation. However, later Qt continued the development of the native desktop GUI components normally, side by side with QML.

Advantages of QML considering Nexstim products include the ability to modify QML's own GUI components to the needs of unique GUI design, make different workflow processes for different applications, brand the GUI for different products, localize GUI for different market areas not only by language but also by layout and possibility to change from Windows platform into Linux.

The software implementation team included a software architect, senior designer, designer, and few consultants. The full software implementation was not completed in the framework of this master's thesis due to a lack of resources. However, the GUI and workflow implementation was adequate for conducting a usability evaluation presented in Section 7.

The user manual and training material were not finished within the time framework of this thesis. The user manual of the existing therapy device is used as a base for creating the manual for the new product. The new user manual is going to be in the form of simple step-by-step instructions for conducting a workflow step. A mock-up of the software integrated user manual is already implemented, but it is incomplete as the content is still missing. The integrated manual is displayed in a web browser styled pop-up window presented in Figure 9. The contextual help is designed to apply the same texts as the manual, but as the content is missing, the implementation of contextual help is postponed.

6.6 User interface evaluation

The manufacturer needs to evaluate the user interface through the development process and after the device is ready to be marketed. The development phase evaluation was known before as usability verification, but with the new version of IEC 62366-1:2015 the name changed into formative evaluation. Moreover, the evaluation of the finished product is now called summative evaluation instead of usability validation.

6.6.1 Formative evaluation

Formative evaluations are performed through the product development process to gain valuable information on the usability of the user interface. At the early sate

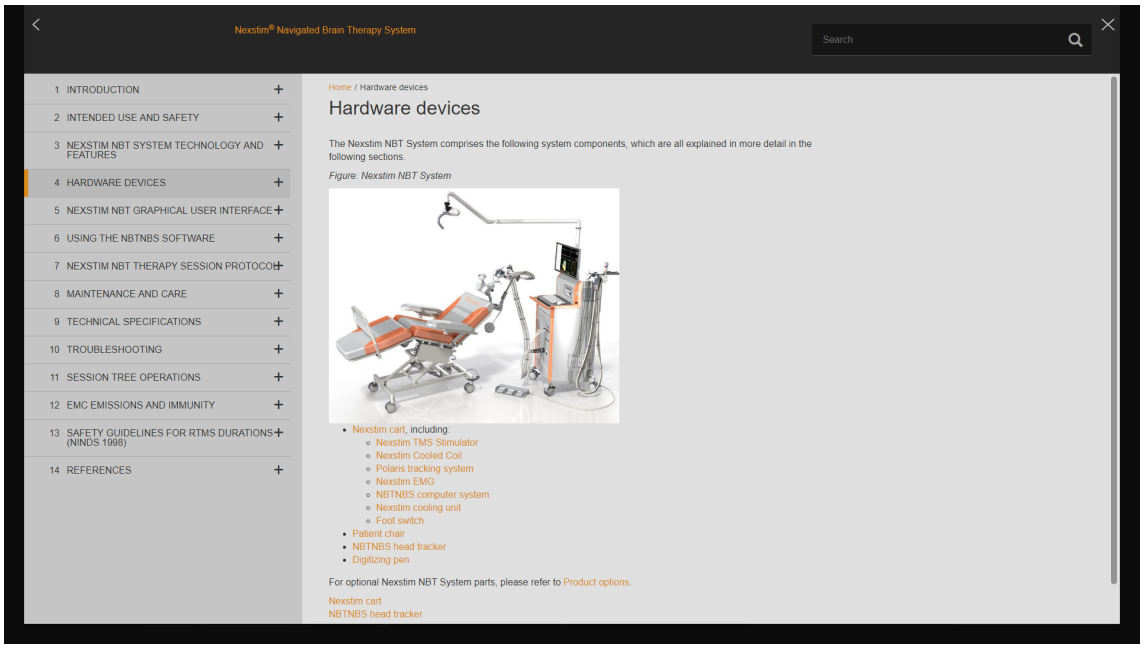
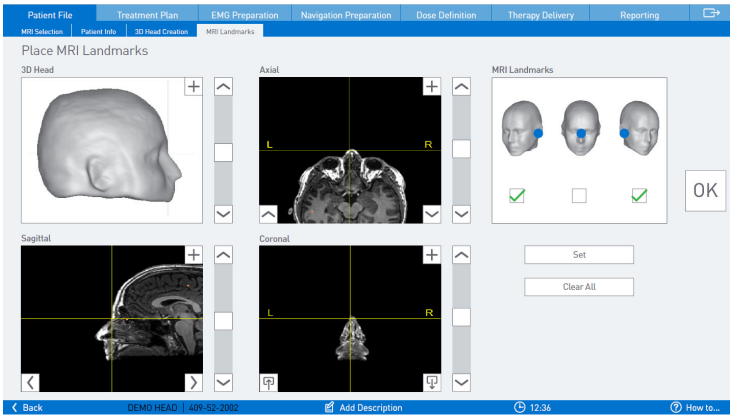
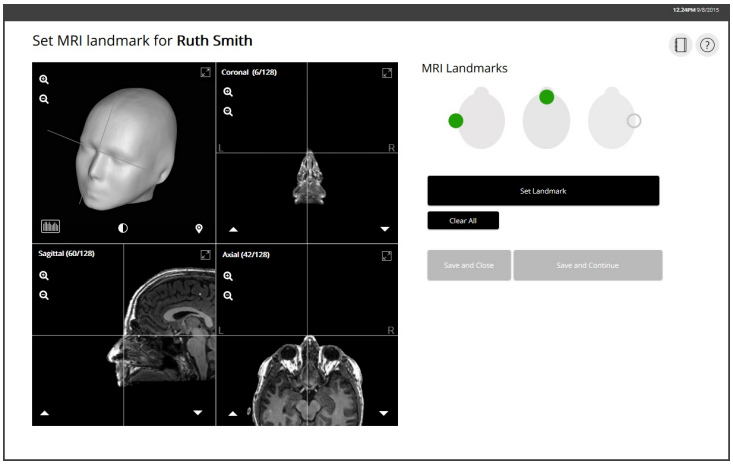


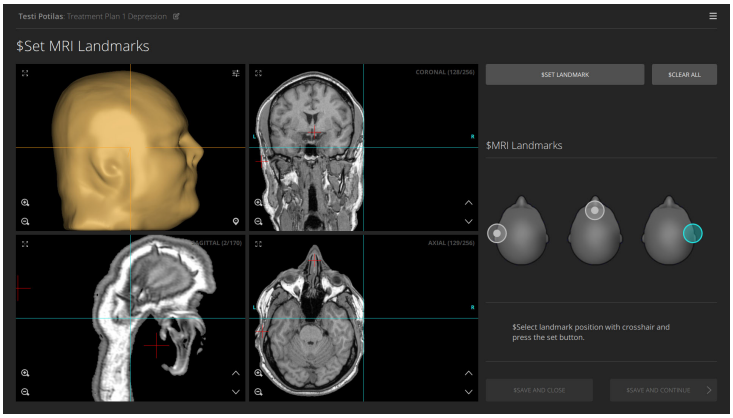
Figure 9: The outlook of the software integrated user manual. The manual content presented in the figure is from the existing NBT device's manual.



(a)



(b)



(c)

Figure 10: Example of screens used in the formative evaluations. All images represent a screen where user sets MRI landmarks to the patient's MR images. a) is an image shown to the participants in the concept test held in February 2015, b) is a screenshot from the UI prototyping software used in the second user test in December 2015, and c) is a screenshot of the latest prototype used in the usability evaluation in 2019.

of development, the formative evaluation is performed to gain information about the design strengths and weaknesses. The assessment conducted at the later state can help the manufacturer to evaluate the safety and usability of the user interface. Formative evaluations can be usability tests, cognitive walkthroughs, expert reviews and other evaluations that will benefit the development of the product.

Nexstim has conducted two user tests with the SNBT user interface. In February 2015, the initial concept of the SNBT software was tested to receive feedback on the preliminary workflow-driven graphical user interface. The study included 11 participants from six facilities that had from seven months to four years of experience with Nexstim's systems. Concept drawings of the GUI were shown and participants were asked to rate the different aspects of the screen designs. One example of a concept drawing is presented in Figure 10a.

The user interface of the concept test was developed for a stroke application and not for the depression application which is now the intended use for the SNBT software. However, the basic principles of the workflow are the same and thus the results can be used as input for the current depression and pain applications. The concept test showed that the workflow orientation and touch screen were convenient and beneficial changes from the user interface of the existing device.

The second user test was conducted in December 2015 using a UI prototyping software. The prototyping software is a convenient and quick tool for developing and testing GUI ideas. At this point in the development process, the appearance of the software design had evolved from the previous as the company designing the GUI changed. An example of the GUI design used in this evaluation is presented in Figure 10b. Six participants from three clinical sites completed seven scenarios with the prototype software from a laptop without the presence of a TMS system. The scenarios included multiple real-world tasks that are frequently done in the treatment protocol. The participant's interaction with the application was monitored by the facilitator and note-takers. Audio and on-screen activity were recorded and keystrokes were tracked and quantified. After completing the tasks, the participants answered to a questionnaire and were interviewed to get additional feedback.

In the data analysis, successful scenario completion ratio, error-free ratio, and subjective evaluations were analyzed. In three scenarios out of seven, all participants were able to complete tasks successfully without help. However, two scenarios resulted in critical errors and none of the participants were able to complete them. The error-free ratio was 100% only in two scenarios and all other scenarios had errors with all participants. In the error-free ratio calculations, all errors were recorded even though the participant might have recovered from the error by themselves. These kinds of errors do not necessarily prevent the completion of the scenario but are frustrating for the user.

While completing the scenarios, participants expressed over two hundred comments about the software. The findings were classified as problems or items that generated use errors, additional findings impacting functionality, graphics-related suggestions and positive comments. The findings included, for example, users not knowing what to do on a screen, failing to locate a certain icon, not understanding the function of a control and so on. Participants particularly liked 25 features.

The overall impression was that the new software is user-friendly, streamlined, well-organized, and intuitive. The participants thought that the workflow guided process may decrease the learning curve for future users. Users particularly liked the idea of controlling the application with a touchscreen. Participants appreciated that they are guided through all necessary steps in the treatment process. Both of the early user tests gave important feedback from the new GUI and software, and the results have been exploited in the further development of the product during the past years.

In 2019, a new user test was conducted for a more advanced user interface. The methods and results of this study are presented in Section 7.

6.6.2 Summative evaluation

The objective of summative evaluation is to validate the usability of the UI of the finished product. The assessment is conducted to confirm that the UI enables effective use and minimizes harmful use errors. The summative evaluation will be performed for the new product when the software is fully implemented and user manual and training material have been established.

The summative evaluation will be conducted by performing a usability test where users are observed while performing tasks using the device. The usability test was selected as the test method because it is quite easy to conduct, and an expert review would not be thorough enough to test such a major change in the GUI. Moreover, expert reviews should not be used when hazard-related use scenarios can arise during use.

The summative evaluation will be conducted under simulated conditions. The use environment will be a simulated clinical room and the test will include healthy subjects acting as patients. The participants will be clinical personnel that are used to using nTMS devices for the treatment of depression and pain.

Summative evaluations have previously been conducted for Nexstim's existing products and thus the focus of the assessment will be to evaluate the usability of the new GUI and new hardware components, the touch screen and the foot pedal. The evaluation will include testing both application and the hazard-related use scenarios recognized in the risk analysis presented in Section 3. The usability test will also assess if the training material and user manual are perceivable and understandable.

The data collected from the usability test will include performance data and subjective comments. Performance data will be recorded by observing users conducting tasks on the new device and by recording correct use, use errors, close calls and use difficulties. The test environment should resemble normal clinical setting as closely as possible and thus participants should not be asked to think aloud and test personnel should not make contact with participants.

The subjective data will be collected by interviewing participants after the usability test. The aim of the interview will be to find out the root causes of any occurred use errors. The interview questions will also cover the overall impression of the new product and any possible close calls and use difficulties that were not recorded during use.

In the data analysis, potential causes of each use error will be analyzed and all root causes will be determined. If new use errors, hazards or hazard-related use scenarios are discovered during the summative evaluation or further development is otherwise needed, the UE process will be re-entered. If no improvements are needed, the residual risks of the device will be evaluated according to the risk management standard ISO 14971.

7 Usability evaluation

The formative evaluation of the new nTMS product was conducted at the end of the year 2019 using a functional prototype. The product had evolved significantly from the last usability studies and thus a new evaluation took place. The aim of this assessment was to verify the usability of the GUI and gain input for further development if usability deficiencies were found. The objective was to study whether the usability goals described in Section 2.4 were met.

7.1 Methods

The usability study evaluated the advanced SNBT prototype using six common scenarios that were completed by current Nexstim system end-users. The participants were asked to carry out the most frequent tasks in the depression treatment protocol as well as troubleshoot some common problematic situations in the environment representing the actual use conditions.

7.1.1 Test system

The test system was created so that the participating hospital's existing NBT system stayed as intact as possible to avoid any technical problems after the usability test. The test system included a laptop with the latest software version of the new product, a touch screen, and a foot pedal. The participating hospital's NBT system patient chair, camera, digitizer pen, and stimulation coils were used. The test room's NBT system display was replaced with the touch screen that was plugged into the laptop and a power source. The NBT system's camera wire was detached and an additional camera wire was connected between the camera and the laptop. A 3D printed head was used to simulate patient as the prototype could not be used with real patients yet.

As the study did not include real patients present, TMS and EMG simulators were used. The TMS simulator creates artificial magnetic pulses that evoke responses to the EMG simulator. The SNBT software communicates with the simulators and thus pulses and responses can be displayed on the screen. The TMS and EMG simulators, foot pedal, RFID reader for head trackers, and touch screen USB were connected to a USB hub on the laptop. The 3D printed patient head was fixed to the patient chair neck rest for the length of the test. A camera stand with an attached smartphone was placed so that the touch screen activities were visible on video. The test system is presented in Figure 11.

The software version used in the usability test was not fully and had some defects. All screens in the workflow were not implemented and thus they were passed by with the help of the test personnel.

7.1.2 Participants

Nine participants from two hospitals participated in the usability study. Participants were from the user (system operator) and admin user groups described in Section

Figure 11: The test setup used in the usability evaluation.

6.1.4. Participants included seven nurses, one physician, and one hospital physicist. Only one physician, instead of the desired two, was able to attend the study due to time constraints. The use experience with Nexstim systems varied between 5 months and 15 years. The participant backgrounds are presented in Table 4. All participants received a short training of the new product and test system before the test started. The quick training was not as extensive as the clinical training given to the new users.

Participant	Occupation	Experience with Nexstim systems
P1	Nurse	1 year
P2	Nurse	8 years
P3	Physician	15 years
P4	Hospital Physicist	4 years
P5	Nurse	4 years
P6	Nurse	4 years
P7	Nurse	4 years
P8	Nurse	2 years
P9	Nurse	5 months

Table 4: Participant information.

7.1.3 Use scenarios

In the usability test, the participants operated the system in six typical use scenarios presented in Table 5. All scenarios included two to eight tasks that were selected because they were the most frequent functions performed in the treatment process or tasks for solving common problematic situations. The completion of selected tasks also represented the fulfillment of some of the usability goals. The tasks were selected from the depression application as it was more familiar for the participants and the software implementation was more advanced than in the pain application.

Participants received a task sheet presented in Appendix A including all scenarios and tasks to guide them through the test. The task sheet was provided both in English and in Finnish. The participants were instructed to read the whole scenario and list of tasks and ask questions, if they had any, before starting the tasks. While performing the tasks, participants were instructed to talk aloud. In addition to the task sheet, the test facilitator asked additional questions that were used to test some of the usability goals.

In Scenario 1, Create patient, participants needed to conduct eight tasks in order to create a new patient. Among others, the scenario included tasks for importing MR images to the system, entering patient information and setting MRI landmarks. Scenario 2 asked participants to prepare the system before the patient arrives for their first baseline visit. In Scenario 3, the participants needed to prepare the patient for treatment by activating a head tracker and aligning the camera so that the tracker was in the camera's field of view. Registering the patient's head and the MR images to the system were also performed in Scenario 3. The patient's first baseline session was conducted in Scenario 4. The baseline session included mapping the patient's motor cortex in order to find the representative area of thumb muscles, determining the stimulation intensity for the treatment, and determining the location where the treatment would be given. Scenario 5 included preparing the system for a patient's 3rd treatment session and finally, in Scenario 6, participants gave treatment for the patient.

Due to time constraints, all participants were not able to complete all scenarios. Participant 7 completed only scenarios three to eight and participant 8 scenarios two to eight. Scenarios 1 and 2 were selected to be left out as creating new patients was not the participating nurses' duty. The fifth participant could not perform tasks 7 and 8 from the fourth scenario due to technical difficulties with the software. This scenario also included two additional tasks given by the facilitator. Users needed to tell the value of the E-field maximum and the EMG response amplitude and describe where the user was in the workflow process.

7.1.4 Data collection methods

Data was collected by video recording the participants' actions on the touch screen while performing the scenarios. This method was selected as it would have been too difficult to write notes from all user actions as there was only one test personnel present. During the test, the users were asked to think aloud to enable the recording of comments and difficulties using the system. After the test, participants filled in a

	Scenario
1	Create patient
2	Prepare for baseline visit
3	Prepare patient and register
4	Conduct baseline
5	Load patient
6	Conduct treatment

Table 5: Use scenarios performed in the usability evaluation.

questionnaire and were interviewed. The questionnaire and interview questions were translated into Finnish to help with answering.

7.1.5 Data analysis

Data analysis aimed at the quantitative evaluation of the user interface and the identification of usability problems and safety issues created by the user interface. The video recordings were analyzed to find out scenario completion ratio, number and type of errors encountered during each scenario and error-free ratio. Scenario competition ratio is the proportion of participants able to complete a scenario. Each scenario required the participant to obtain or input specific data. The scenario was considered completed when the user had finished the last task of a scenario, indicated that the goal of the scenario was obtained, or the participant requested and received guidance from test personnel.

The errors in conducting scenarios were classified into three types. An error was considered as critical if it produced an incorrect outcome or restrained participant from continuing the treatment process. Asking for help from the test personnel in the essential tasks for the treatment process was also seen as a critical error. Another class of critical errors was incorrect outcomes or asking help in the troubleshooting tasks that were not mandatory in the treatment process. These faults were called moderate critical errors (MCE). The third class of misuse was non-critical errors (NCE) that were deviations from the optimal task completion and were recovered by the participant, or if they were not detected, they did not result in processing problems or unexpected results. Questions about the user interface, unrelated to helping the participant through a task, were considered as comments. Forgetting to do a troubleshooting task was not considered as an error as they do not block the workflow process. Some NCEs and minor problems that did not significantly affect the task completion were marked as difficulties. A difficulty was, for example, extra clicks on the screen and wandering in the GUI when looking for a function.

Based on the above-mentioned error classification and the frequency participants made those errors, a severity classification was made for each problem the participants encountered during the usability test. The highest class of severity was Class 4 in which the usability problems generated possible risk for the system user or the patient. Class 3 severity was assigned to critical errors as they prevented the participant from correctly completing the task. Moderate critical errors and non-critical errors,

made by over 25% of participants, were in severity Class 2. The lowest class of severity, Class 1, was assigned to NCE and MCEs that were made by less than 25% of participants and difficulties experienced by over 25%. The severity classification is presented in Table 6.

Severity	Explanation
Class 4	Usability problems that generate possible risk for the user or patient.
Class 3	Critical errors in essential tasks in the treatment process that prevent correct use.
Class 2	MCEs and NCEs made by over 25% of participants
Class 1	NCEs and MCEs made by less than 25% of participants and difficulties experienced by over 25% of participants.

Table 6: Severity classification for problems encountered during the usability test.

The post-task questionnaire and interview answers were analyzed qualitatively to find out the users' opinions and features participants generally liked and disliked.

7.2 Results

7.2.1 Scenario completion

Scenario 1, where participants needed to import MR images and create a new patient, was performed only by seven participants due to time constraints. From those seven participants, four were able to complete the scenario successfully without critical errors and none completed the scenario error-free. The scenario completion and error-free ratios for all scenarios are presented in Table 7. Critical errors were faced in task 1. Import MR images, task 3. Optimize MRI and a troubleshooting task 7. Set point of interest (POI). Further analysis of these problems is presented in Section 7.2.2 Identified usability findings. Scenario 1 resulted in seven critical errors, 6 moderate critical errors, 15 non-critical errors and 6 difficulties.

In scenario 2, participants needed to open previously created patient data and prepare the system for the patient's first baseline visit. All eight participants who performed the scenario were able to complete the scenario successfully and error-free.

The usability of the GUI when preparing the patient for the baseline was tested in scenario 3. Eight out of nine participants were able to perform this scenario successfully and error-free. The critical error was made when one participant didn't understand what align camera meant. The participant making this error was a physician that seldom performs this step in the treatment process.

The most difficult scenario was the fourth one, where participants needed to map the hand area, determine MT, and create a target for therapy. The scenario completion ratio was only 2/9 and error-free ratio 1/9. The biggest problem was that the participant did not know what to do in each task. Task 4, perform preliminary

coil orientation map, created the most errors as participants thought they were supposed to determine the motor threshold in this step. Five out of six participants, who remembered to do task 3, were able to find the coil expiration date from the GUI quite easily. The scenario included two additional tasks given by the test personnel that were not in the task sheet. All users were able to see the value of the E-field maximum and EMG response amplitude when standing next to the patient. The test personnel forgot to ask the participants to perform the second additional task. The task would have been to describe where in the workflow process the user was. In total, scenario 4 yielded in 17 critical errors, one MCE, 15 NCEs, and 8 difficulties.

The fifth scenario, where participants needed to prepare the system for patients 3rd treatment, had only one critical error and one non-critical error. The completion ratio was 8/9 and error-free ratio 7/9.

In scenario 6, participants gave treatment for a patient. Both the completion ratio and error-free ratio were 4/9. However, only one error was made in the essential tasks of the treatment process. A participant tried to pause the treatment by clicking different parts of the GUI causing a non-critical error. Troubleshooting task 3, adjust camera orientation with the help of the alignment tool, yielded the most errors as the participants could not find the alignment tool. In total, scenario 6 yielded in 0 critical errors, 5 MCE, 4 NCEs and 4 difficulties.

	Scenario	Completion ratio	Error-free ratio
1	Create Patient	4/7	0
2	Prepare for baseline visit	8/8	8/8
3	Prepare patient and register	8/9	8/9
4	Conduct baseline	2/9	1/9
5	Load patient	8/9	7/9
6	Conduct treatment	4/9	4/9

Table 7: Completion ratio and error-free ratio for each scenario.

7.2.2 Identified usability findings

The data analysis resulted in the discovery of a total of 17 usability findings presented in Appendix B. Most problems arose from legibility issues regarding instructions in the workflow step screens and pop-up windows and from the accessibility of functions. None of the identified deficiencies were classified as the severity Class 4 implying possible risk for the user or patient.

Six usability findings were given Class 3 severity as they created critical errors that prevented correct use in the treatment process. One of these problems was created in scenario 4 task 1 where participants needed to map the patient's motor cortex. The screen header, "Preliminary Mapping", was not informative enough for the participants to know what to do. The test personnel needed to explain the task which created a critical error.

In severity class 2, five usability findings were identified. The problems were found as participants tried to perform tasks in incorrect screens or pop-up windows

and they could not find functions. These problems were faced by over 25% of the participants. For example, users tried to set a POI in a pop-up where the user can change the 3D head and MRI scan properties. The two windows were said to be too similar and thus users mixed them up. Participants also had problems finding the tools for aligning the camera and viewing information about the used stimulating coil.

Less frequent problems were grouped into severity Class 1. These six deficiencies were created by a lack of training and ambiguous features in GUI design. When reviewing maps, participants had difficulties selecting a stimulus from a table. The table had different colors for rows that were highlighted and selected. The participants thought that a row was selected by simply highlighting it and not by selecting the radio button. Figure 12 presents the map review screen with one highlighted row.

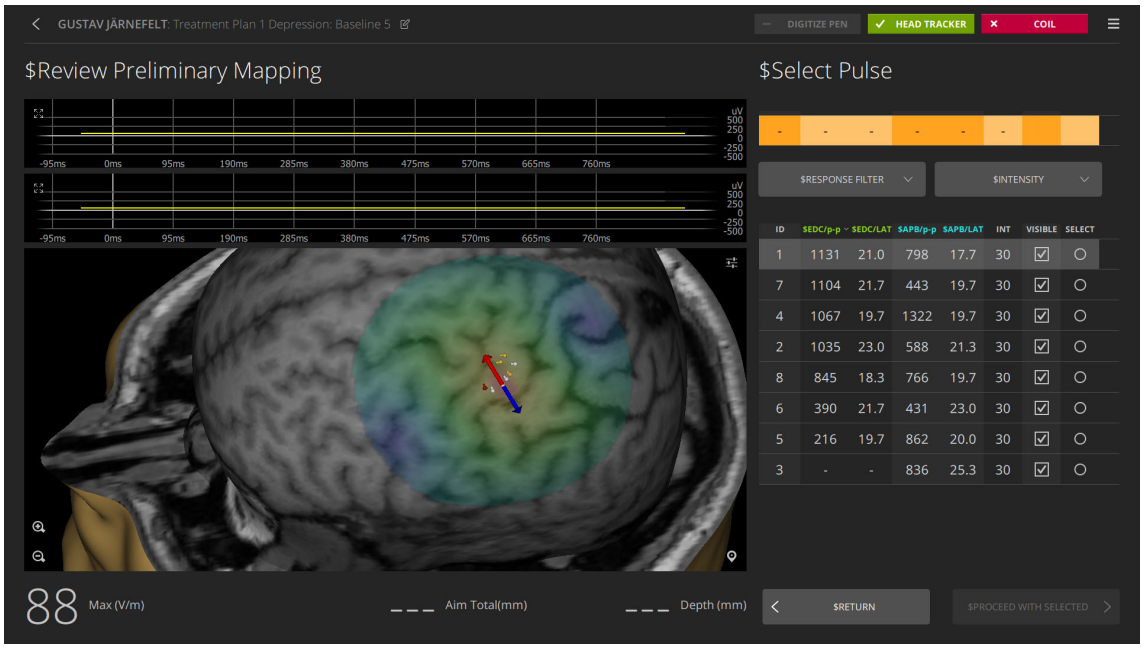


Figure 12: Review Mapping screen with first stimulus highlighted. User can select a stimulus by tapping the select radio button on the right side of the stimulus table.

A GUI design deficiency was found as one participant wanted to set her own POI instead of using the DLPFC depression therapy POI that the software had suggested. The user could not find the tab where she could have created her own POI. The POI pop-up with the DLPFC-tab open is presented in Figure 13. The POI tab could have been opened by tapping the dim orange text link, "POI", above the row describing the DLPFC POI information.

7.2.3 Questionnaire and interview

The questionnaire included a question and 10 claims that participants needed to rate after using the system. The question asked how likely the participant would recommend the software used today to their colleagues on a scale from 0 to 10, 0

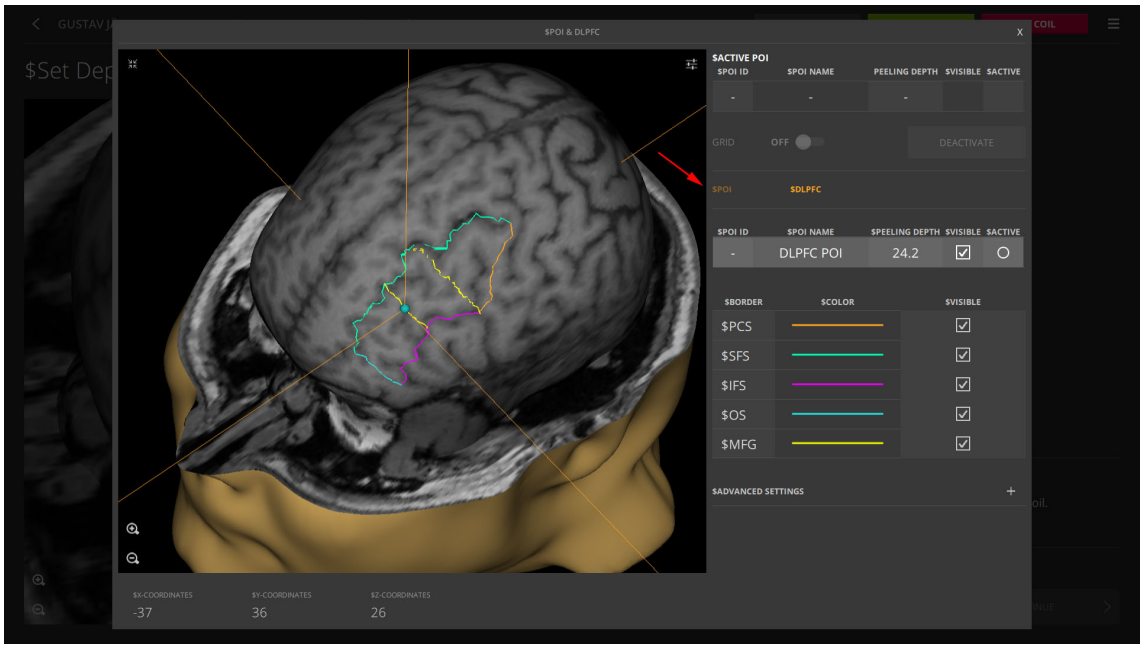


Figure 13: POI pop-up window with DLPFC-POI-tab open. User can open POI-tab by tapping the dim orange text "POI" indicated with a red arrow.

meaning "highly unlikely" and 10 meaning "extremely likely". The results show that participants would recommend the SNBT software to their colleagues very likely with an average score of 9.

The questionnaire claims and answers are presented in Table 8. The results show that the software was well-liked. Participants strongly agreed that they would enjoy the frequent use of the software and that most people would learn to use it quickly. Users agreed that the product is easy to use, and that various functions were well integrated. Participants disagreed with the claims that the software is complex, inconsistent, cumbersome to use and that users would need to learn a lot before using the software independently. The largest variance in the questionnaire resulted from the claim about needing technical support when using the system. Three participants strongly disagreed with the claim and three somewhat agreed. Users didn't feel confident nor insecure when using the system

In the interview, participants said that the GUI is modern and more pleasant than the NBT software they are using now. Users particularly liked that the software guides them through the treatment process. The touchscreen, ability to move the 3D head with gestures and the new pedal that can be used with bigger shoes received praise.

Most users said that it is hard to detect any annoying features during such a so short use. Two participants thought the sounds in the registration were annoying. The sound when a landmark is set should be clearer and there should not be a loud cue after the 3-point registration as the registration process is still not ready. Few participants said that they had problems finding functions in the new software as they looked for the same icons and buttons as in the software they are using now.

Usability Claims	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
1. I think I would like to use this software frequently.	-	-	-	3	6
2. I found the software unnecessarily complex.	2	5	1	1	-
3. I thought the software was easy to use.	-	-	1	5	3
4. I think that I would need the support of a technical person to be able to use this system.	3	1	2	3	-
5. I found the various functions in this software were well integrated.	-	-	2	5	2
6. I thought there was too much inconsistency in this software.	1	7	1	-	-
7. I would imagine that most people would learn to use this software very quickly.	-	-	-	4	5
8. I found the software very cumbersome to use.	4	5	-	-	-
9. I felt confident using the software.	-	1	6	1	1
10. I need to learn a lot of things before I could get going with this software.	4	4	1	-	-

Table 8: Questionnaire scores. Each row presents a claim and number of participants answering the claim according to the column header.

However, the participants said that they would perform better if they had a chance to do the test again as they already learned a lot when using the system for the first time.

Some terminology used in the software were considered hard to understand. The participating nurses said that most difficult terms were in the parts of the workflow where a physician would usually be present and thus nurses would not need to understand them. It was noted that some screen headers and button texts could be more informative.

Participants emphasized that the software is very easy to use and that they would learn to use it quickly. Most of the participants thought that they could train a colleague after using the system a few times to treat patients. When asking would the participants like to have their possible new device to be Nexstim's current NBT

or the new SNBT, seven out of nine told they would like to have SNBT. One user was unsure, and one explained that they would probably need to have a more flexible system for their research purposes.

7.2.4 Meeting the usability goals

One aim of the formative evaluation was to study whether the usability goals were reached. Result analysis discovered that three out of six goals were met. Goal 2 was reached as users were able to use the system from two distances. Participants could hear and see information standing next to the patient as well as standing close to the screen. Goal 4 was met as all users were able to describe the treatment quality and accuracy in the treatment summary screen. Goal 6, less downtime, was met as participants were able to find information about the coil expiration date.

Goal 1 was not met as users needed help on what to do next. This goal can be met by better wording in the screens and by adding training. Goal 5, it is clear to the user where one is in the treatment process, was not met as the test personnel forgot to ask the users to describe the screen they were in in scenario 4. Goal number 3, ability to conduct treatment session in less than one hour, was not tested in the usability test as fulfilling this goal requires more experience with the new system and this was the first time the users operated the system. Testing goal number 3 would also have needed real patients present.

8 Discussion

The regulation of medical devices is a truly complicated field and mastering it requires a lot of effort and experience. The legal text of regulations is hard to comprehend and most rules have exceptions and repealing rules for special cases. This master's thesis described only the outline of medical device regulation in the target market areas of the new device. The regulation and device approval processes have been considered to be more strict and less efficient in the US than in Europe. Now, as new requirements have been added to the MDR, it will be interesting to see if the approval of new devices will become more difficult and take more time.

The usability engineering process was developed for the company to fulfill the requirements of the IEC 62366-1:2015 usability standard. The standard operating procedure (SOP) for the process was approved by the head of QA&RA and taken as part of the company's quality system. The usability engineering process was applied and required product development activities and documents were delivered. The process was proven to be adequate in analyzing the usability of the new product. Especially task analysis was found to be an excellent method for analyzing possible safety-related use error. Task analysis generated new use errors that had not been found in risk analysis of the company's previous devices.

The developed UE process SOP is a document containing a list of deliverable documents and requirements for their content. The SOP could be further developed by adding specific instructions on how the process was conducted for the new device and specifying what methods have proven to be suitable for the company. For example, task analysis and severity assessment of usability findings in the formative evaluation could be added to the SOP.

The usability evaluation gave important information about the new product and the usability of the GUI. In general, the device received great feedback and participants enjoyed using the system. Most tasks the participants performed during the evaluation were completed successfully. It was determined that the most frequent tasks in the treatment process were easier to conduct than the tasks that participants seldom did.

The evaluation resulted in some usability findings that were mostly due to bad wording choices in the screens and pop-up windows and lack of user training. Some findings were made about functions being difficult to find due to ambiguous GUI components. Problems using the system might also have aroused from the fact that it was the very first time the participants used the system and they tried to find similar icons and buttons as in the existing system they had used before. The test results may have been better if the participants had had an earlier contact with the system. Many participants thought that they could perform better if they had the chance to do the test again.

Most of the usability findings could be handled by rephrasing the screen headers and instruction texts more carefully. The headers should have a clear description of what the user should do in the screen. The training given before use should emphasize the new GUI components and differences compared to the existing devices the participants have used. The GUI design for the tabs was found confusing and a

redesign should be considered.

The test setup was practical for the usability evaluation and could easily be regenerated for future tests. The software version used was adequate enough for the assessment although some software defects contributed to the use errors. Furthermore, a more advanced software version would certainly give more representative usability data. One great improvement would be the implementation of the EMG and relevant graphs. Interpretation of EMG responses is an important factor during the motor cortex mapping and the usability of graphs should be carefully studied.

The test participant profile was quite homogeneous as seven out of the nine participants were nurses. It would have given more realistic data if more physicians and hospital physicists would have participated as some duties in the treatment process are typically assigned only for them. Prior experience using Nexstim's systems varied greatly among participants, but there were no significant differences in task completion between different backgrounds.

The scenarios and tasks that were selected for the usability evaluation were the ones that are normally done in the treatment process. The task descriptions in the task sheet could have been more informative. The descriptions were kept succinct to not give too much direction for completing the tasks. Another and possibly superior solution compared to giving more instructions would be to reword the screen headers as previously suggested.

Video recording screen activities while participants performed the tasks was discovered to be an appropriate data gathering method. Going through and tracking all use errors and comments from the over five hours of video recording was a time-consuming but more precise method than writing notes while observing participants performing the tasks would have been. Quantitative analysis of usability is quite difficult, but the selected methods were deemed satisfactory. Severity classification for usability findings was a great method for assessing further developments regarding the GUI components.

In the future, a new usability evaluation should be performed with a more advanced prototype where the usability findings discovered in this thesis have been minimized. The assessment should be performed in the US as it is the main target area for the new product. The training given before the test should be more extensive and focus on certain new GUI components and differences compared to the existing devices. As all usability goals were not met in the usability evaluation presented in this thesis, the usability goals should be retested. The test scenarios should include testing the pain application as the workflow is more flexible than in depression application and therefore it might be harder to use.

The summative evaluation presented in Section 6.6.2 should be performed when the new device is ready for production. If the final usability assessment is successful and no new hazard-related use scenarios are discovered, the usability of the device is proven to be adequate. However, the evaluation of the product does not end when the device is released to the marked as it needs to be continuously evaluated in post-market surveillance.

9 Summary

The number of medical devices has increased rapidly over the years and they have become an essential part of health care. The safety and effectiveness of medical devices are ensured by laws and regulations that are overseen by authorities such as the Food and Drug Administration (FDA) in the United States (US).

In 2017, the European Parliament and Council published a new medical device regulation (MDR) that repealed the old medical device directives. Considering usability, the regulation states that the risks related to use error should be eliminated or reduced as far as possible. Complex user interfaces (UI) that have inadequate usability can cause use errors that can possibly lead to dangerous situations.

The goal of this master's thesis was to research medical device regulation, update the quality system of a company to the level of the new MDR, develop an usability engineering (UE) process according to usability standards and apply the process for a new navigated transcranial magnetic stimulation (nTMS) device.

The regulation of medical devices was discovered to be a complicated field as all countries have their own laws and regulations that differ from each other. In all studied regulations, medical devices are divided into risk-based classes and the required proof of the safety and effectiveness varies between classes. The manufacturer of the device needs to demonstrate safety before the product can be released to the market.

This master's thesis renewed the quality management system (QMS) of an nTMS device manufacturing company to comply with the usability requirements of the MDR. The QMS was updated by compiling a standard operating procedure (SOP) for conducting the UE process described in the harmonized standard IEC 62366-1:2015. The UE process was applied according to the requirements of the usability standard.

The UE process resulted in finding several use errors and hazards that could affect the safety of the patients or the users. The hazard-related use errors were further analyzed in risk analysis and the severity of each risk was assessed. Previous formative evaluations were described briefly and an outline for the final summative evaluation was presented.

The usability evaluation of the new product conducted at the end of the year 2019 was presented more thoroughly. The goal of the assessment was to verify the usability of the graphical user interface (GUI) and test the fulfillment of usability goals. In the study, nine current nTMS device users operated the new system in the most frequent tasks in treatment in a realistic clinical use environment. The data analysis identified usability deficiencies by tracking the scenario completion ratio, the number and type of use errors, and the error-free ratio. The severity of the usability findings was assessed with a four-point scale taking into account the type and frequency of the use errors.

The evaluation resulted in the identification of 17 usability findings that were mostly created by unclear instruction texts and ambiguous GUI components. None of the findings generated possible harm for the user or the patient but less severe findings were discovered. Three out of the six usability goals were met.

The developed UE process was successful, but the process could not be finished

as the product did not get finished in the time frame of this thesis. In the future, the UE process should be continued by performing a new formative evaluation with a further developed system and finalized by conducting a summative evaluation for the production equivalent system.

References

[1] World Health Organization. Medical Devices: Managing the Mismatch: an outcome of the priority medical devices project: methodology briefing paper, 2010.

[2] Rados C. Medical Device & Radiological Health Regulations Come of Age. https://www.fda.gov/about-fda/histories-product-regulation/medical-device-radiological-health-regulations-come-age, 2006. [Online: accessed 26-January-2020].

[3] Sorenson C. and Drummond M. Improving Medical Device Regulation: The United States and Europe in Perspective. *Milbank Quarterly*, 92(1):114—-150, 2014.

[4] Council of European Union. Council Directive of 20 June 1990 on the approximation of the laws of the Member States relating to active implantable medical devices (90/385/EEC), 1990. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:01990L0385-20071011&from=EN.

[5] Council of European Union. Council Directive 93/42/EEC of 14 June 1993 concerning medical devices, 1993. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1993L0042:20071011:EN:PDF.

[6] Council of European Union. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices., 1998. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1998L0079:20090807:EN:PDF.

[7] European Parliament. REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. *Official Journal of the European Union*, L 117, 2017.

[8] IEC 62366-1:2015. Medical devices – Part 1: Application of usability engineering to medical devices. Standard, International Electrotechnical Commission, Geneva, Switzerland, 2015.

[9] Barker A. T., Jalinous R., and Freeston I. L. Non-invasive magnetic stimulation of human motor cortex. *The Lancet*, 1:1106–1107, 1985.

[10] Krieg S. M. *Navigated Transcranial Magnetic Stimulation in Neurosurgery*. International series of monographs on physics. Springer International Publishing, 2017.

[11] Ruohonen J. *Transcranial magnetic stimulation: Modelling and new techniques*. PhD thesis, Espoo: Helsinki University of Technology, 1998.

[12] Fava M. Diagnosis and definition of treatment-resistant depression. *Biological Psychiatry*, 53(8):649–659, 2003.

[13] Baxter L. R., Schwartz J. M., and Phelps M. E. Reduction of Prefrontal Cortex Glucose Metabolism Common to Three Types of Depression. *Archives of General Psychiatry*, 46(3):243—250, 1989.

[14] Martinot J. L., Hardy P., Feline A., and Huret J. D. Left prefrontal glucose hypometabolism in the depressed state: a confirmation. *American Journal of Psychiatry*, 147(3):1313—1317, 1990.

[15] Lefaucheur J. P., André-Obadia N., Antal A., Ayache S., Baeken C., Benninger D., Cantello R., Cincotta M., de Carvalho M., De Ridder D., Devanne H., Di Lazzaro V., Filipović S., Hummel F., Jääskeläinen S., Kimiskidis V., Koch G., Langguth B., Nyffeler T., Oliviero A., Padberg F., Poulet E., Rossi S., Rossini P., Rothwell J., Schönfeldt-Lecuona C., Siebner H., Slotema C., Stagg C., Valls-Sole J., Ziemann U., Paulus W., and Garcia-Larrea L. Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS). *Clinical Neurophysiology*, 125(11):2150–2206, 2014.

[16] European Parliament and of the Council. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, 2017.
`https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02017R0745-20170505&from=EN`.

[17] European Parliament and of the Council. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU, 2017.
`https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02017R0746-20170505&from=EN`.

[18] ISO 13485:2016. Medical devices – Quality management systems – Requirements for regulatory purposes. Standard, International Organization for Standardization, Geneva, Switzerland, 2016.

[19] Lauer M., Barker J. P., Solano M., and Dublin J. FDA Device Regulation. *Missouri Medicine*, 114(4):283–288, 2017.

[20] FDA. A History of Medical Device Regulation & Oversight in the United States. `https://www.fda.gov/medical-devices/overview-device-regulation/history-medical-device-regulation-oversight-united-states`, 2019. [Online: accessed 19-August-2019].

[21] Kramer D. B., Shuai X., and Kesselheim A. S. Regulation of Medical Devices in the United States and European Union. *The new England journal of medicine*, 399(9):848–855, 2012.

[22] Jarow J. P. and Baxley J. H. Medical devices: US medical device regulation. *Urologic Oncology*, 33:128–132, 2015.

[23] QUALITY SYSTEM REGULATION. Code of Federal Regulations. `https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=820`, title 21 § 820.30(a)(2) 2019. [Online: accessed 30-January-2020].

[24] FDA. Regulatory controls. `https://www.fda.gov/medical-devices/overview-device-regulation/regulatory-controls#special`, 2018. [Online: accessed 19-August-2019].

[25] FDA. Premarket Approval (PMA). `https://www.fda.gov/medical-devices/premarket-submissions/premarket-approval-pma`, 2019. [Online: accessed 19-August-2019].

[26] FDA. Evaluation of Automatic Class III Designation (De Novo). `https://www.fda.gov/medical-devices/premarket-submissions/evaluation-automatic-class-iii-designation-de-novo`, 2019. [Online: accessed 19-August-2019].

[27] Health Canada. Legislation and Guidelines - Medical devices. `https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/legislation-guidelines.html`, 2011. [Online: accessed 9-September-2019].

[28] Ståhlberg T. *Terveydenhuollon laitteiden lakisääteiset määräykset kansainvälisillä markkinoilla Suomi ja EU fokuksessa.* Tekes, 2015.

[29] Health Canada. Safe Medical Devices in Canada. `https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/dhp-mps/alt_formats/pdf/md-im/activit/fs-fi/meddevfs_matmedfd-eng.pdf`, 2014. [Online: accessed 9-September-2019].

[30] Lamph S. Regulation of medical devices outside the European Union. *Journal of the Royal Society of Medicine*, 105(1 suppl):12–21, 2012.

[31] World Health Organization. WHO global model regulatory framework for medical devices including in vitro diagnostic medical devices, 2017.

[32] Winters J. M. and Story M. F. *Medical Instrumentation: Accessibility and Usability Considerations.* CRC Press, 2019.

[33] IEC 60601-1-6:2013. Medical electrical equipment – Part 1-6: General requirements for basic safety and essential performance – Collateral standard: Usability. Standard, International Electrotechnical Commission, Geneva, Switzerland, 2013.

[34] FDA. Recognized Consensus Standards. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfStandards/search.cfm, 2018. [Online: accessed 2-January-2020].

[35] European Commission. Commission communication in the framework of the implementation of the Council Directive 93/ 42/EEC concerning medical devices. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=uriserv:OJ.C_.2017.389.01.0029.01.ENG, 2017. [Online: accessed 2-January-2020].

[36] Health Canada. List of Recognized Standards for Medical Devices. https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/standards/list-recognized-standards-medical-devices-guidance.html, 2019. [Online: accessed 2-January-2020].

[37] Rossi S., Hallett M., Rossini P. M., Pascual-Leone A., and The Safety of TMS Consensus Group. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, 120(12):2008–2039, 2009.

[38] Wassermann E. M. Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the international workshop on the safety of repetitive transcranial magnetic stimulation, june 5–7, 1996. *Electroencephalography and Clinical Neurophysiology*, 108(1):1–16, 1997.

[39] EN ISO 14971:2012. Medical devices - Application of risk management to medical devices. Standard, European Committee for Standardization and European Committee For Electrotechnical Standardization, Brussels, Belgium, 2012.

A Use scenarios and tasks

The Table below presents the task sheet given to the participants in the usability test. Tasks marked with an asterisk (*) are additional troubleshooting tasks that are not mandatory in the workflow of treatment.

Scenario	Task
1. You have a new patient coming tomorrow. Please make preparations as far as makes sense today. You have the name of the patient and a USB drive with patient's MR images. Patient name: *Gustaf Järnefelt*.	1. Import MRI 2. Create Patient 3. Optimize MRI 4. Enter Patient Infiromation 5. Select depression application 6. Set MRI landmarks 7.* Set POI to hand knot (POI = point of interest, known as therapy target in NBT) 8. Go to home screen (Save and Quit)
2. The patient you prepared yesterday is coming in 10 min or less. Start preparing for the first baseline visit. Patient name: *Gustaf Järnefelt*	1. Open existing patient 2. Open treatment plan created earlier 3. Add new baseline 4. (End at EMG page)
3. Patient Gustaf Järnefelt arrives for his baseline mapping session. Starting at the EMG page, work through the screens as if you are setting up the disposables, and complete both landmark and advanced registration. The scenario is completed once you tell me your registration accuracy	(Start at EMG screen) 1. Activate Head Tracker 2. Align camera 3. Perform registration 4. Tell the registration accuracy
4. Patient Gustaf Järnefelt has arrived for a mapping session. You have already set up the patient and completed registration. Starting at the Advanced Registration page, complete a mapping session. The scenario is completed once you can view your treatment target on the Depression target summary screen.	(Start at advanced registration screen) 1. Perform preliminary mapping (give max 15 pulses) 2. Review the map and select the best pulse 3.* Check the expiration date of the coil 4. Perform preliminary coil orientation map 5. Review the map and select the best pulse 6. Determine MT 7. Set Depression target 8. (End at Depression target summary screen)

5. Patient *Mikko Mallikas* is coming in 10 min or less for their 3rd treatment session. Start at the Home screen and end once you have gone as far as possible without *Mikko Mallikas* present.	1. Locate existing patient *Mikko Mallikas* 2. Select Start 3rd treatment session
6. Patient arrives for their 3rd treatment session. You have already set up the patient and completed registration. Work through a treatment session starting from the advanced registration page.	(Start at advanced registration) 1. Continue to therapy 2. Select rTMS file for the treatment and continue with the previous MT 3.* Adjust camera orientation with the help of the alignment tool 4. Start therapy (let the therapy run for a minute) 5. Pause therapy 6. Stop therapy 7.* Describe how accurate the therapy was 8. Go to home screen

Table A1: Task sheet given to the participants in the usability evaluation.

B Usability findings

The table below presents the findings discovered in the usability evaluation divided for each severity level. The task column presents the scenario (S) and the task (T) the finding was detected in.

Severity level 1				
None				
Severity level 2				
#	**Task**	**Finding**	**Supportive evidence**	**Note**
1.	S4 T1	The users did not understand what to do in preliminary mapping	The users did not know if they had to use the stimulation coil, the users asked help	
2.	S4 T4	The users assumed that they needed to determine MT although they were in map coil orientation screen	The users attached the coil to the holder and walked to the system in order to start the MT determination, the users had to be instructed to rotate the coil and give stimuli	The coil orientation map has been removed from the application
3.	S1 T3	The users did not know what optimize means	The users asked what they were supposed to do	
4.	S5 T2	A user did not find "START 3RD TREATMENT" button	The user tapped other components on the screen and the test personnel needed to instruct the user	Only one participant faced this problem
5.	S1 T1	The users did not find their way to import MRI because the USB flash drive was not connected	The users went to "Export MRI" screen and later to Patient Work	The users who had connected USB went straight to import MRI
6.	S3 T2	A user did not understand what align camera meant	The test personnel needed to instruct participant to move the camera to the target	The problem was faced only by one participant who usually does not prepare the patient for the treatment
Severity level 3				
	Task	**Finding**	**Supportive evidence**	**Note**

7.	S4 T2	The user thought they could select pulse already on the mapping screen	The users reviewed and highlighted pulses already in the mapping screen	Software defect: highlighting pulses should not be available in the mapping screens
8.	S1 T7	The user thought they could set POI in adjust overlay	The users opened the adjust overlay to peel the 3D head, moved the crosshair to the location they wanted to set the POI to, and pressed "APPLY" button. However, the POI was not set as the users were in the wrong pop-up.	
9.	S6 T3	The users did not find the align camera overlay	5/9 of the participants needed help to find the overlay	
10.	S4 T3	The users did not find the coil information overlay	The users tried to find the overlay in the Menu	Only one participant asked help
11.	S1 T8	The users did not go to the home screen	The users tapped “SAVE AND CONTINUE” button instead of ”SAVE AND CLOSE”	
Severity level 4				
	Task	**Finding**	**Supportive evidence**	**Note**
12.	S4 T7	A user was not able to find the POI tab in the POI overlay when the DLPFC-POI tab was open	The user asked help	Only one participant tried to change tabs
13.	S4 T7	The users thought therapy target was given by setting a DLPFC-POI and not by using the coil	The users wandered in the "Set target" screen not sure how to continue	
14.	S1 T6	The browsing of the MRI slices was difficult	The users changed the slices one by one by tapping arrows in MRI slice view	

15.	S1 T6	The users had difficulties selecting a stimulus	The users thought highlight was already select and the users wandered in the GUI trying to find the select radio button.	
16.	S1 T7	The users had difficulties finding the "CREATE POI" button.	The user wandered in the GUI trying to find a way to create the POI	
17.	S4 T2	A user had difficulties finding the pause button	The user tapped several places in the GUI to pause the treatment before finding pause button	Only one participant faced this problem

Table B1: Usability findings discovered during usability evaluation presented in Section 7.